

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 15.1684 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481B-1
Perfect score: 5956
Sequence: 1 MAARVLLLTALTLCHGFNLD.....FKRQYKDMSEGPPGAEPQ 1153

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5956	100.0	1153	1 RWHULB	cell surface glyco
2	4545	76.3	1153	2 S00551	leukocyte surface
3	3513	59.0	1163	1 RWHULC	cell surface glyco
4	1559	26.2	1170	2 S03308	cell surface glyco
5	1549	26.0	1163	2 I56126	lymphocyte fuction
6	1163	19.5	1179	2 A53213	integrin alpha-1 c
7	1108.5	18.6	1151	2 A45226	integrin alpha-2 s
8	1099	18.5	1170	2 I45914	integrin alpha 2 s
9	1089	18.3	1178	2 S44142	VLA-2 protein homo
10	1087.5	18.3	1181	2 A33998	integrin alpha-2 c
11	1082	18.2	1180	2 A35854	integrin alpha-1 c
12	667	11.2	1039	2 A41131	lymphocyte-Peyer's
13	638	10.7	1038	2 S06046	integrin alpha-4 c
14	630	10.6	1035	2 I58409	integrin alpha-9 c
15	614.5	10.3	1041	2 T31437	integrin alpha cha
16	579.5	9.7	1054	2 JC7294	alphan integrin -
17	572.5	9.6	1051	2 A35761	cell surface glyco
18	567.5	9.5	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.3	1053	2 A44250	integrin alpha-5 c
20	550	9.2	1034	2 A36108	integrin alpha-v c
21	535	9.0	1044	2 T10050	integrin alpha-v c
22	532	8.9	1049	2 A27079	fibronectin recept
23	532	8.9	1073	2 B36429	integrin alpha-6 c
24	530.5	8.9	1072	2 A38457	integrin alpha-6 c
25	529.5	8.9	1051	2 A40021	integrin alpha-3 c
26	526	8.8	1048	2 A27421	integrin alpha-5 c
27	525.5	8.8	1091	2 A41543	integrin alpha-6 c
28	517	8.7	1044	2 S16516	integrin alpha-8 c
29	505	8.5	1394	2 A29637	position-specific

30 496.5 8.3 1146 2 S40311 integrin - fruit f
31 495.5 8.3 1039 2 A34269 integrin alpha-2b
32 492.5 8.3 1037 2 A60163 glycoprotein I1b -
33 490 8.2 1137 2 JCS950 integrin alpha-7 c
34 488 8.2 1135 2 I61186 alpha-7 integrin -
35 486 8.2 126 2 B30892 leukocyte adhesion
36 474.5 8.0 1226 2 S44824 PS4F2.1 protein -
37 469.5 7.9 1106 2 S38783 integrin alpha cha
38 454.5 7.6 1139 2 S28277 hypochetical prote
39 454 7.6 1045 2 S60571 integrin alpha v c
40 430.5 7.2 1115 2 T09433 integrin alpha cha
41 423.5 7.1 1115 2 T09403 integrin alpha cha
42 391 6.6 764 2 I36916 glycoprotein I1b -
43 309 5.2 1086 2 T18523 integrin alpha cha
44 304.5 5.1 272 2 A55348 integrin alpha-1 -
45 299 5.0 604 2 I36917 glycoprotein I1b -

ALIGNMENTS

RESULT 1

RWHULB

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3,
B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <OR>

A:Cross-references: GB:J03925; NID:gl87284; PIDN:AA59544.1; PID:g307148

A:Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AA5A

A:Note: the authors translated the codon TAC for residue 1129 as Thr

A:Note: part of this sequence, including the amino end of the mature protein, was conf
R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally r

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:gl80018; PIDN:AA58410.1; PID:g553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe
A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept
A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:Cross-references: GB:J04145; NID:gl89068; PIDN:AA59903.1; PID:g386975

A:Note: part of this sequence was confirmed by protein sequencing
 R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 480-490, 1993
 A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
 n during evolution.
 A:Reference number: A46526
 A:Accession: A46526
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499,501-1153 <FLE>
 A:Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
 A:Note: the last three bases of intron 13, CAG, are included in some but not all mature
 A:Note: sequence extracted from NCBI backbone (NCBIP:121963)
 R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A:Reference number: A90664; MUID:87076671; PMID:3539202
 A:Accession: A26091
 A:Molecule type: protein
 A:Residues: 17-31 <PIE>
 A:Experimental source: granulocytes
 R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A:Title: Characterization of the myeloid-specific CD11b promoter.
 A:Reference number: I52567; MUID:92144986; PMID:1346576
 A:Accession: I52567
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-9 <RES>
 A:Cross-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219
 C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
 C:Genetics:
 A:Gene: GDB:ITGAM; CR3A
 A:Cross-references: GDB:120599; OMIM:120980
 A:Map position: 16p11.2-16p11.2
 A:Note: promoter contains a GATA motif and two Sp1 consensus binding sites
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
 F:17-1108/Domain: extracellular #status predicted <EXT>
 F:148-318/Domain: von Willebrand factor type A repeat homology <VMA2>
 F:465-473/Region: calcium/magnesium binding #status predicted
 F:530-538/Region: calcium/magnesium binding #status predicted
 F:593-601/Region: calcium/magnesium binding #status predicted
 F:1109-1134/Domain: transmembrane #status predicted <TM>
 F:1135-1153/Domain: intracellular #status predicted <INT>
 F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 100.0%; Score 5956; DB 1; Length 1153;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MALRVLLLTALTLCCHGNLDENAMTFQENARGFGSVVQLOGSRVVGAPQBIIVANQR 60
 DB 1 MALRVLLLTALTLCCHGNLDENAMTFQENARGFGSVVQLOGSRVVGAPQBIIVANQR 60
 QY 61 GSLYQCDYSTGSCPIRLQVPVEAVNMSGLSLAATSPQLLACGPTVHQTCSNTYVK 120
 DB 61 GSLYQCDYSTGSCPIRLQVPVEAVNMSGLSLAATSPQLLACGPTVHQTCSNTYVK 120
 QY 121 GLCFLFGSNLRQOPKPFPEALRCQPEDSDIAFLDGSGLIIPHDFRMKFEFVSTVMEQL 180
 DB 121 GLCFLFGSNLRQOPKPFPEALRCQPEDSDIAFLDGSGLIIPHDFRMKFEFVSTVMEQL 180
 QY 181 KSKTFLFSLMQYSEERIFHTTFKFNQNPNSRLVLPITOLLGRTHATGIRKVVRELFN 240
 DB 181 KSKTFLFSLMQYSEERIFHTTFKFNQNPNSRLVLPITOLLGRTHATGIRKVVRELFN 240
 QY 241 ITNGARKNAPKILVITDGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQEL 300
 DB 241 ITNGARKNAPKILVITDGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQEL 300

RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N;Alternate names: complement-3 receptor alpha chain

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999

C;Accession: S00551; I59078

R;Pytela, R.

EMBO J. 7, 1371-1378, 1988

A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the

A;Reference number: S00551; MUID:88312584; PMID:3044779

A;Accession: S00551

Query Match	Best Local Similarity	Score	3513; DB 1; Length 1163;
Matches	701; Conservative	138; Mismatches 296; Indels	6; Gaps 4;
QY	5	VLLLTALTCHGFNLDTENAMTFQENARGFCQSVVLOGSRVNVGAPQEIIVAAHQSGSLY	64
Db	8	LLLFATALSLGFLNDTEELTAFRVDSAGFGDSVVQVYANVWVVGAPQKITAANQQTGGLY	67
QY	65	QCDYSTGSCPEIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENVTYKGLCF	124
Db	68	QCYSTGACEPIGLQVPEAVNMSLGLSLASTTSPQLLACGPTVHCEGRNMYLTGLCF	127
QY	125	LFGSNLRQPOKPFPEARLGCPOEDSDIAFLDGGSGSIIPHDPRRMKEFVSTVMEQLKSK	184
Db	128	LLGPT--QLTQRLPVSRQECPRQEQDIVFLDGGSGSISSRNFAFMNMFRAVISQFORPS	185
QY	185	TLFSLMQVSEBFRTHFTKFEQNNPNRSLVKPTTOLLGRTHTATGIRKVVRELFNIYNG	244
Db	186	TQFSLMQFSNKFQTHFTPEEBERTSNPLSLASVHLQGLQFTYTATAIQNVVHRLFHASYG	245
QY	245	ARKNAFKILVITDGEKFGDPLGYEDVTPEDADREGVIRYVGVGDAPFSEKSRQELNTIA	304
Db	246	ARRDATKILVITDCKEGDSLVDYKDVI PMADAAGIIRIYATGVGLAFQNRNWSKELNDIA	305
QY	305	SKPPRDHVQVNFPEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGP	364
Db	306	SKPQSEHIFKVEDFDALKDQIQNLREKIFAIEGTETTTSSSFELEMAQEGFSAVFTPDGP	365
QY	365	LLSTVGSYDMAGGVFLYTSKSKSTPFIMNTRVDSMDNDAYLGVAIAIIIRNRVQSLVLGAP	424
Db	366	VLGAVGSPWTSGGAFLYPPNMSPTPIINMSQENVDNRDSYLGYSTELALWKGVSILVLGAP	425
QY	425	RYOHIGLVAMFRQTMKESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYE	484
Db	426	RYOHTGKAVITQVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTLVLVIGAPHYYE	485
QY	485	QTRGGQVSCVCLPRQCARWOCDAVLYEQGQPWGRFGAALTVLGDVNGDKLTDVAIGAP	544
Db	486	QTRGGQVSCVCLPRGWR--RWNCDAVLYEQGHPGRFGAALTVLGDVNGDKLTDVVGAP	544
QY	545	GEEDNRGAVYLFHGTSGSGI SPFSHSQR IAGSKLSPLQYFGQSLGGQDLTMDGLVDLTV	604
Db	545	GEENRGAVYLFHGVLGSPISPSHSOR IAGSQLSSRLQYFGQALSGGQDLTDQGLVDLAV	604
QY	605	GAQGHVLLRSQPLRVKAI MEFNPREVARNVFECDNQVKGKEAGEVRVCLHVOKSTRD	664
Db	605	GARGQVLLRTPVLWVGVSQWQIFPAEIPRSFAFCEQVVSQTLVQSNICILYIDKRSKN	664
QY	665	RLREGQISVTVYDLALDSGRPHSAVFNETKGNSTRROTQVLGLTOTCETLKLQLPNCIE	724
Db	665	LLGSRDLQSSVTLDALDPGRLS PRATQETKNRSLSRVRVLGLKAHCENFNLLPSCVE	724
QY	725	DPVSPIVLUNFSLVGTTPLSAFGNLRPVLAEADAQRALTALPFPEKNCNGNDINTCODDLSIT	784
Db	725	DSVTPITLURLNFTLVGKPLAFPLNRPLMLADAQRYFTASLPFEKNCGADHICQDNLGIS	784
QY	785	FSPMSLDCLVGCGPREFNVTVTVRNDGEDSYRTQVTFPFLDLSYRKVYSTLQNRQSRWS	844
Db	785	FSFPGKSLVGSNLELNAEVMVMNDGEDSYRTTITTFSHPAGLSRYRYAEGQKQGLRSL	844
QY	845	RLACESASSTVEGALKSTPSCSINHP IIPENSEVTFNITFDVDSKASLGNKLLKANVTS	904
Db	845	HLTCDSAPVG--SQGTWSTNSCRINHLIFGGGAQITFLATFDVSPKAVLGDRLULLTANVSS	902
QY	905	ENNMPRTNKTFQLELPVKIYAVMVVVTSHGVSTKYLNETAS--ENTSRVNMHQYQVSNLQO	963

[illegible]

Db 133 YTTGVCSDVPDF-QURTSFAPAVQTCF-SFIDVVVVCDSESNIIYPWD--AVKNLEKXV 188
QY 178 EOLK--KSKTSLFMOYSEEFRIHFTFKFQNNPNSLVKPIITOLL-----GRTHATGCI 231
Db 189 QGLDIGPTKQGLQIYANPRVFMNTFKSKD---EMIKATSTQFYGGDLTWTFKAI 245
QY 232 RKVRELFNITNGARKNAFKILVITDGBKFGDPLGYEDVPEADREGVIRVIGV----287
Db 246 QYARDTAYSTAAGRGPCATKVMVVVTDGESH-DGSKLKAVIDQCNDNLRFGIAVLGYL 304
QY 288 -GDARFSEKSRQELNITIASKPRDRHVQVNNFEALKTIONQREKIFALEGTEGTGSSSF 346
Db 305 NRNALDTKNLKEIKAINSIPTERHFNVSDEADLEKAGTIGEOIFSIEGTVOG-GDNF 363
QY 347 EHEMSQEGFSAAIT--SNGPLSLTVSGYDWAGVFLYTSKESSTFINMT--RVDSDMN-D 401
Db 364 QHEMSQVGSFAEYSPQNNILMGLGAYDWSGTIVQKTHGHLIFSKQAFQILQDRNHS 423
QY 402 AYLGYAAAAILNRVOSLVGLAPRYQHIGLVAMFRONTQGWESNANV-----KGTQIGA 455
Db 424 SYLGVSASISTGNSVHFVAGAPRANYTGQVLYSVN-----ENGNTVTIQSGRQDQIGS 478
QY 456 YFGASLCSVDVDSNGSTDVLICAPHYEQTR--GGQVSVCLPQROARWQCDVLYGE 513
Db 479 YFGSVLCAVDVWVKDTITDVLVAGAPMYMNDLKEEGRVYLFITTKG-ILNWH--QPLEG 535
QY 514 QGQPMCRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGIGSPSHSORIA 573
Db 536 NGLNARFSGAIALSDIINWGDNDVIVGSPLENQNSGAVIYNGHEGM-IRLYSQKIL 594
QY 574 GS--KLSPRIQYFGOSLSGGQDITMDGLVLTGAQGHVLLRLRSQVPLRVKAIMFNPRE 631
Db 595 GSDRAFSSHLQYFGSLDYGDLNGDSITDVSVGAFGQVQVLMWSQISADVSDASPTPK 654
QY 632 VARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVTDLALD-----SGRH 687
Db 655 I--TLNKNABE-----KLKLCF---SAKFRPTNQNQVAIVNITIDEDQFSRVI 701
QY 688 SRAVFNETKNSRROTVOGLTQCE--TLKQLPNCIEDPVSPVILRLNFSL--VGTP 743
Db 702 SRGLFKENNERCLOKTMVISOAQCSEYIIHIEPS---DIISPLNLCMNISLENFGT--756
QY 744 SAFGNLRPLVAEDAQRLEFALPFPEKNCNDNITCDDLSITF---SPMSLCLVVGGR 799
Db 757 -----NPALAEASETVKVFSPFPHKDCGDGVICISDLVNLVQOLPATQOQPIVSNQ 810
QY 800 EFNVTVRNMGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESAST-EVSG 858
Db 811 RLTFVQLKNKKEASVNTETVVDVSENLF-----ASWMPVDGTEVTQOIAS 858
QY 859 ALKSTCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTQFOL 918
Db 859 SQKSVTCNVGYPALKSKQQTFTTFNFDNLQ-NLQOQASIFRALSESQENMADNSVL 917
QY 919 ELPKYAVYVMTVSHGYSTKYLNTASENTSRYMHQYQVSNLQOR-----SLPI 968
Db 918 KLSLYDAEIHIT-RSTNINFEVSLDGNVSVV-HSFE--DIGPKFISIKVTGSGVPV 973
QY 969 SLVFLVRLNQTVWRDPQVTFSEN--LSSTCHTKE-----RLPSHS 1009
Db 974 SMA-----SVIIHPIQVTKDKNPLMYLTGVHTDQAGDISCEABINPLKIGQTSVV 1024
QY 1010 DELAE-LRKAPVNCSTAVCORIQCDIPFFGIQEEFNATLKGNSLFDWYIKTSHNHLIV 1068
Db 1025 SFKSENPRHIELKNCRTASCNIMCWLDIQVGEYFLNVSTRIWNGTFAASTFTQVOLT 1084
QY 1069 STAEI-LFNDVSFTL-----LPQOGAFVRSQTETKVEPPE-VFNPLIVGSSVGLLL 1120
Db 1085 AAABIDYNYQIYIEENTVITP-----LTINKPHEKVEVPTGVIVGSVIAGILL 1134
QY 1121 LALITAALYKLGFPKQYKDM 1141
Db 1135 LALVAIWLKGLGFPKRYKYM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Lar

submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but no

A:Reference number: S44142

A:Accession: S44142

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A:Cross-references: EMBL:Z29987; NID:9473098; PID:CAA82877.1; PID:9473099

C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F:169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.3%; Score 1089; DB 2; Length 1178;

Best Local Similarity 28.1%; Pred. No. 6.5e-67;

Matches 348; Conservative 209; Mismatches 492; Indels 188; Gaps 44;

QY 5 VLLLTALTLCGHFNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIIVAANQR 60

Db 15 LMLVQGIILNCLAYNVGLFGAKIFSGPSEQFGYSVQQLTNPQGNLWLVGSPWSPENRM 74

QY 61 GSLXQC--DYSTGSCPEIRLQ-----VPVEAVNMSLGLSLAATTSPPQLLACGPTVHQ 111

Db 75 GDVYKCPVDLPTATCEKUNLQNSASISNVTETIKTNMSLGLTLTRNPGTGGFLTCGLPLAH 134

QY 112 TCSNTYVKGCLFLFGSLNRLOQPO---KPEALRGCPQEDSDIAFLIDGSGSIIIPHFR 168

Db 135 QCGNQYATGIC---SDVSPDFQFLTSPFAVQACPSL-VDVVVVCDSESNIIY-WEA 187

QY 169 MKEFVSTVMEQK--KSKTLFSLMOYSEFRIHFTFKFQNNPNSLVKPIITQLG-RT 225

Db 188 VKNFLVKVFTGLDIGPKKTOVALIOYANEPRIIFNLNDFETKEDVMQATSETROHGGDLT 247

QY 226 HTATGIRKVVRELFNITNGARKNAFKILVITDGBKFGDPLGYEDVPEADREGVIRVI 285

Db 248 NTFRAIEFARDYAYQTSQGRPCATKVMVVVTDGESH-DGSKLKTVIOCNDDLEIRFGI 306

QY 286 GV-----GDARFSEKSRQELNITIASKPRDRHVQVNNFEALKTIONQREKIFAIEGTQT 340

Db 307 AVLGYLNRNALDTKNLKEIKAIASPTERYEFNVADAEALLKAGTLGEOIFSIEGTQV 366

QY 341 GSSSFHEHMSQEGFSA--AITSNGPLSLTVSGYDWAGVFLYTSKESSTFINMT--RVD 396

Db 367 G-GDNFQMEMAQVGSADYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIFPKQAFDQVL 425

QY 397 SDMN-DAYLYGAAAILNRVQSLVGLAPRYQHIGLVAMFRONTQGWESNANV---KGT 451

Db 426 QDRNHSSFLUGYSVAALISTEDGVHVFAGAPRANYTGQVLYSVNK---QGNVTVIOSHRGD 482

QY 452 QICAYFGASLCSVDVDSNGSTDVLIGAPHYEQTR--GGQVSVCLPQROARWQCDVAV 509

Db 483 QIGSYFGSVLCSVDVDKDTITDVLVAGAPTMYNDLKEEGKVYLFITTKGILLNQH---F 539

QY 510 LYGEQOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGISPSHS 569

Db 540 LSGPECTGNARFGSAIALSDINMDGFNDVIVGSPVENENSGAVIYNGHQT-IRTKYS 598

QY 570 QRTAGSKLSPR--LQYFGOSLGGODLTDMDGLVLTGAQGHVLLRLRSQVPLRVKAIMBF 627

Db 599 QKILSGNAGFRHLHOFFGRLSDGYDLNGDSITDVSIGALGQVIQWSSQSIADVALEALF 658

QY 628 NPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVTDLALD---S 683

Db 659 TP-----DKITLLNKDAKITLKLCFRAEFPAQONNQV--AILFNMTLADAGHS 705

QY 684 GRPHSAVFNETKNSTRTOTVGLTQTCET--LKLQLPNCIEDPVSPVILRLNFSLVGT 741

Db 706 SRVTSRGVFNSESRFQKNNVNVNKKSEHHISQKPS---DVNPLDLRVDISLENP 762
Qy 742 PLSAFGLNRPVLAEDAKRLFTALPFPEKNCNDNICODDLSI-----TFSPMSLDC 792
Db 763 GTS-----PALEYSETVKVSPFPFYKEGSDGICISDLILDVQQLPAQTQSF----- 811
Qy 793 LVGGPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLACSSAS 852
Db 812 IVSNQKRLTFPSVLKNGRESAYNTVLAEPSENLF-----ASFSPVVDGTE 859
Qy 853 ST-BVSGALKSTCSINHPPIPEENSEVTFTNTPVDVSKASIGNKLLKLLKANVTSENNWPT 911
Db 860 VTCEVGSQKSVTCVGVFALKSQVTFTFINFDNLQ-NLQNAANFQAFSESQ--ET 916
Qy 912 NKTE--FOLELPKAVYVMVVTSHGVSTKYLNFTASENTSRVMQHQVQVSNLQR----- 964
Db 917 NKADNSVSLTPLYDAELHLT-RSTNINFEIISDENAPSVIK---SVEDIGKPFPSL 972
Qy 965 -----SLPISLVFLV-----PVRLNQTVIWD 985
Db 973 KVTAGSAPVSMALVTIHPQVTEKNPLLYLTGTQDQAGDISCTAEINPLKLPHTA--- 1029
Qy 986 RPQVTF-SENUSSTCHYKRLPSHSDFLAELRKAPVNCSTAVCQRICQDIPFGIQEFP 1044
Db 1030 -PSVSFKENPR---HTKE-----LDCRTSGSNITCWLKDLHMAEY 1068
Qy 1045 NATLKGNSLPDIWYIKTSHNHLIIVSTAEILFNSDVFTLLPQGAFFVRSQTEKVEPEVP 1104
Db 1069 FINVTTRWNTFAASTFQTVOLTAABEIDTHNQLFVIBENAVTPIPMIMKPTKEAEP 1128
Qy 1105 NPLPLIVGSSVGGLLALITAAALYKLGFFKRYQKDM 1141
Db 1129 T-GVIIGSIAGILLALLTAGLWKLGLFFKRYQKDM 1163

RESULT 10

A33998
Integrin alpha-2 chain precursor - human
N:Alternate names: Cd49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 chain
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 Sequence revision 18-Sep-1992 #text_change 21-Jul-2000
C:Accession: A33998; B56793; A53117
J: Takada, Y.; Hemler, M.E.
R: Cell Biol. 109, 397-407, 1989
A: Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A: Reference number: A33998; MUID: 89308879; PMID: 2545729
A: Accession: A33998
A: Molecule type: mRNA
A: Residues: 1-1181 <TAK>
A: Cross-references: GB: X17033; NID: G33906; PIDN: CAA34894.1; PID: G33907
A: Note: The authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R: Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A: Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIc*, GPIIa and GPIIb) from the human alpha-2 integrin gene promoter. Identification of positive and negative regulatory regions.
A: Reference number: A53117; MUID: 94103255; PMID: 8276836
A: Accession: A53117
A: Molecule type: DNA
A: Residues: 1-16, 'V', 18-21 <ZUT>
A: Cross-references: GB: L24121; NID: G400342; PIDN: AAA16619.2; PID: G4583535
A: Note: authors translated the codon GTA for residue 17 as Leu
C: Genetics:
A: Gene: GDB: ITGA2; CD49B
A: Cross-references: GDB: 128031; OMIM: 192974
A: Map position: 5q11.1-5q11.2

C: Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C: Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F: 1-29/Domain: signal sequence #status predicted <SIG>
F: 30-1133/Domain: extracellular #status predicted <EXT>
F: 172-347/Domain: von Willebrand factor type A repeat homology <WMA2>
F: 1134-1154/Domain: transmembrane #status predicted <TM>
F: 1155-1181/Domain: intracellular #status predicted <CYT>
F: 105, 112, 343, 432, 460, 475, 699, 1057, 1074, 1081/Binding site: carbohydrate (Asn) (covalent)

Query Match

18.3%; Score 1087.5; DB 2; Length 1181;

Best Local Similarity 27.3%; Pred. No. 8 3e-67;

Matches 339; Conservative 214; Mismatches 498; Indels 193; Gaps 44;

Qy 3 LRVLTLTALT-----LCHGNLDTENAMTFQ-ENARGFGOSVVQL-----QSGRVVVGAPQE 53

Db 11 LPLLLVLALSGQILNCCLAYNGLPEAKIFSGSPSEQGYAVQOQFNPKNMILLVGPWS 70

Qy 54 IVAANGSLYQC--DYSTGSCPEIRLQ-----VPEAVNMSLGLSLAATTSPPOLLA 104

Db 71 GFPENRMGVDYKCFVDLSTATCEKLNLTQSTSPNVTETMKNMISLGLILTRNMGTTGGFLT 130

Qy 105 CGPTVHTCSENTVWVKGCLFGLFGLNLRQOPKQFPEALRGCPQEDSDIAFLIDGSGSIIPH 164

Db 131 CGPLWAOQGNQYTTGVCSDISDPDF-QLSASESPATQCPSL-IDVVVCDSESNISYPW 188

Qy 165 DFRMKFEFVTVMEQLK--KSKTLFSLMYSEBEFRHFTPEFQNNPNRSLVKPITQLL 222

Db 189 D--AVKNFLEKFGVGLDIPGTTKTOVGLIQYANNRVVFNLTNYTKKEMIVATVSQTQSG 246

Qy 223 G-RTHATGTRKVVRELNFNTNGARKNAFKILVITDGEKFGDPLGVEDVIEPADREGVI 281

Db 247 GDLTNTFGAIQYARKYASAGRRSATKVMVVTGDESH-DGSMUKLAVIDQCNHNL 305

Qy 282 RYVIGV-----GDAFRSEKSRQELNTIASKPRDHFQVNVFALKTIONLREKIFAIE 336

Db 306 RFLAVLGVNLRNALDKNLIKETKATASIPTEFFNFVDEAALLKAGTLGQIFISIE 365

Qy 337 GTGTGSSSFEHMSQEGFSAITSNGP--LLSTVGSYDMAGGVFLYTSKEKSTFINMT- 393

Db 366 GTVQG-GDNFQEMSQVGFSDYSSQNDILMLGAVGAFGWSGITVOKTSHGLIFPKQAF 424

Qy 394 -RVSDMN-DAYLGIAAAAILRNVRVSLVIGARQYHQHIGLVAMPQNTGMWESNANV--- 448

Db 425 DQILQDRNHSYLGYSVAAISTGESTHFVACAPRANTGQIVLYSVN-----ENGNTVI 479

Qy 449 ---KGTQIGAYFGASLCSVDVDSNGSDTLVIGAPHYVEQTR--GGOVSVCPPLRGORAR 503

Db 480 QAHRGDDQIGSYFGSLCSVDVDKDTITDVLVGLAPMYMSDLKKEEGRVYLFITIKKGLGQ 539

Qy 504 WQCDVLYGQGPWGRFGAALTVLGDNVNGDKLTDAVAGPGEEDNRGAVVLPFGTSGSG 563

Db 540 HQ---FLEGPEGIENTFPGSAIAALSINMDGFNDVVGSPLENQNSGAVIYNGHQT- 595

Qy 564 ISPSHSQRIAGS--KLSPLRQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSPQVLRV 621

Db 596 IRTKYSQKILGSDGAFRSHLQYFGSLDGYGLNGDSITDVSIGAGVQVQLWSQSADV 655

Qy 622 KATMEFNPVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRGOQTSVVTYDLAL 681

Db 656 AIEASFTPEKI--TLVNKNAQII-----LKLCF-----SAKFRPTKONNOVAIVYMITL 702

Qy 682 D----SCRPHSAVFNETKNSRROTQVGLGTQTC--ETLKLQPNCIEDPVSPIVLRNLN 735

Db 703 DADGFSRVTSRGLFKENNERCLOKNNVNAQASCPEHIYIQEPS---DVVNSLDLRVD 759

Qy 736 FSLVGTPLSAFGLNRPVLAEDAKRLFTALPFPEKNCNDNICODDLSITF-----SFMSLD 791

Db 760 ISLENPCTS-----PALEYSETKVSIPFFHKDCGEDGLCISDLVDLRQIPAAEQBP 813

Qy 792 CLVVGGRPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLACESA 851

Db 814 FIVSNQNKRLTFSVTLKKNRESAYNTGIVVDSENLFF-----ASFSLPVDGT 861

[illegible]

RESULT 12

A4131
 lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
 N;Alternate names: integrin alpha-4
 C;Species: Mus musculus (house mouse)
 C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999
 C;Accession: A4131; S16742
 F;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmänn, B.; Weissman, I.L.
 J. Cell Biol. 115, 1149-1158, 1991

852	Qy	SST-EVSGALKSTSCSINHPIPPENSEBVTNITFDVDVDSKASLGNKULLKANVTSENMPER	910
862	Db	EVTQVAAASOKSVACDVGYPALKREQVITFDNFNLQ-NLONQASLFOALSESQEEEN	920
911	Qy	TNKTEQLELPVKYAYVMVTVSHGVSTKYLYNFTASENTRVMQHOYOVSNLQOR	964
921	Db	KADNLVNLKIPLLYDAEI---HLTRSTNINFYEISDGNVPSIVHSFEDVGPKFIFSLK	976
965	Qy	---SLPISLIFVLV---PVRNLNQTVIWDR	986
977	Db	VTTGSPVPSMATVIIHPIQYTKERNPLMYLTGVQTDKAGDISCNADINPLKIQOT	1032
987	Qy	PQVTF--SENLSSTCHTKERLPSSDFLAEURKAPVNVNCSTAVCQRIQCDDIPFGIOERN	1045
1033	Db	SSVSFKSENR---HTKE-----LNCRTASCNSVTCMLKDVHMKGEYF	1072
1046	Qy	ATLKGNLSPDWYIKTSHNHLLIVSTABI-LFNDOSVFTLLPGQAGFVRSQETKVBEPFVP	1104
1073	Db	VNVTTIRWNGTFASSTFQTVQLTAAAEINTYNPETIYVI-----EDNTVTIPLMIM	1122
1105	Qy	NP-----LP--LIVGSSVGGLLLLALITALYKLGFPPKROYKDM	1141
1123	Db	KPDEKAEPVGTGIIGSTIAGILLILLALVAITLWKLGFPPKRYEKM	1166

RESULT 11

A35854

integrin alpha-1 chain precursor - rat

C: Species: *Rattus norvegicus* (Norway rat)

C;species:	Rattus norvegicus (Norway rat)
C;Date:	23-Oct-1990
#sequence	revision 13-Sep-1991
#text	change 20-Sep-1999

C;Date: 23-Oct-1990 #Sequence
C:Accession: A35854: S11243

C;Accession: A35854; SII243
P:Ignatius M.J.: Large T.H.: Houde M.: Tawil J.W.: Barton. A.: Esch. F.: Carbonetto,

R; Ignatius, M.J.; Large, T.H.; Houde, M.; Lawli, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol 111 709-720 1990

J. Cell Biol. 111, 709-720, 1990

A;Title: Molecular c
A-Reference number:

A; Reference number: A3

A;Accession: A35854

A;Status: preliminary

A; Molecule type: mRNA

A;Residues: 1-1180 <IGN>

A;Cross-references: GB:X52140; NID:g56493; PIDN:CAA36384.1; PID:g56494

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat

C;Keywords: cell adhesion; cytoskeleton; transmembrane protein

F;170-345/Domain: von W

Query Match	18.2%	Score 1082;	DB 2;	Length 1180;
-------------	-------	-------------	-------	--------------

Best Local Similarity 27.7%; Pred. No. 2e-66;

Matches 351; Conservative 199; Mismatches 484; Indels 234; Gaps 47;

[illegible]

QY 6 LLLTALTCHGNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIVAANQRG 61

[illegible]

Db 18 LLTVILGFCVSFNVDVKNSSMSFSGPVEDMFGYTVQQYENEEGKWVLIGSPLVGQPKARTG 77

THE EFFECT OF THE 1990S ON THE ECONOMIC GROWTH OF THE COUNTRIES OF THE MEDITERRANEAN

62 SLVOCDYSTGSCEP-IRLOVPVEA-----VNMSLGLSLAATTSPOLLACGPTVHQ 111

QY 62 SLIQCDSIGSCEF-IRLQVFEA-----VNISLQDSMAI13FFVTRCOTAVIQ 111

78 DIVKCDVGRERAMPCKI.DI.PVNTSTPNVTEIKENMTEGSTI.-VTNPNGGFLACGPLYAY 136

DB 78 DVYRCFVGRERAMPCKUJLPVNTSIPNVTETKENMTFGSL-VINPNGFLACGFELIAI 130

112 TCSENTVIKCI CEY ECSENI BOOPBOKEFEALBCCBOEENSDTAEI.DGSGSITPHDERRMKE 171


```
Db 648 ALGAVKQISLNISISNLGDDAYDANFVNSRELFF-----INMWQKEENGISCLELLES 701
Qy 854 TEVSGLKSTSCSNHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNK 913
Db 702 DFL-----KCSVGFPMRSKSYEFVFDTSHLG--GEEVLSFIVTAQSG--NTER 750
Qy 914 TE-----FOLELPVKVAVMVVTS-----HGVSTKVLNFTASENTS---RVMQHQY 956
Db 751 SESLHDNTLVLMVPLMHEVDTSITGIMSPTSFVVGESVDAANFIQLDDLECHQPINIL 810
Qy 957 QVSNLQORSPLISLVLFPVRLN-----QTVINDRPOVTFSENLSSTCHTKYER 1004
Db 811 QVNTGPTSTLPGSSVSISFNNRLSSGAEMFHVQEMVVGQKNCSPQKNPTPCIIPOEQ 870
Qy 1005 LPSHSDFLAELRKA-----PVMNCSTAVQCRIQCDIPFFGQIEEFNATLKGNSLFD 1055
Db 871 ENIHTTIFAFPTKGRKVLDCPKIGISCLTAHCN-----FSALAKEESRTI-----D 917
Qy 1056 WYIKTSHNHLIIVTAEILFNDSVFTLLPGQCAFVRQSQTETKVEP----- 1100
Db 918 IY-----MLANT-EILKDDSSVIQ-----FMSRAKVVDPALRVVVEIAHGNPEEV 962
Qy 1101 ---FEVFN---PLPLIVG-----SSVGGLLLLALITAAALYKLGFFKQYKDDMM 1142
Db 963 TVVFEALHNLPRGVVVGWIIAISLLVGLIFLLAVLLVMKQGFRRRYKEII 1015

RESULT 15
Integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:G3220240; PID:G3220241; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.3%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 4.7e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

Qy 352 QEGFSAITNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS-----DNNDAYLG 405
Db 181 QAGFSGIIFSDNSALVMGAPGSYYLQGIYVQSLNRSV-VQATQESNTGTYSFDSNYRG 239
Qy 406 YAAAI--ILNRVQSLVLGAPRYOHI-GLVAMFRONTGMWESNANVKGTOIGAYFGASLC 462
Db 240 YSLALGDFNGGVQDYVVGTPRABSLMGLVAIFDNLNQN-----QVMGTQIVAYFGYSVT 296
Qy 463 SVDVDSNGSTDLVLIGAPHYVEQTRGGQSVSCLPRGORARWQCDVLYLGEQ----- 514
Db 297 VVDI--NNDYDLDLVGAPMYMDGPAIQ-----RWEAGAVVYVLQNPDPVPGA 343
Qy 515 -----GQFWRFGAALTVLGDVNGDKLTVDVAIGAPGEEDNRGAVYLFHGTSGS 562
Db 344 SNRLSSTLTGGQIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGCDAGAVYIYHG--SAN 402
Qy 563 GISPSHQRTAGSKLS-PRLOYFGQSLGGODLTMGDLVLTCAQ--GHVLLRSQPVL 619
Db 403 GLKSTPAQVLTPTLGHSGITTFGFSLQGGQMDKNKYPDLLVGAESANTAVLIRTPV 462
Qy 620 RVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHLVQKSTRDLRREGIQSV----- 674
```

```
Db 463 SLDATLNTSP-----IGINLENKTYE-LADGTMTVTSFIAMT 497
Qy 675 -----VTYDLALDSG-RPHSRAVNETKNSRRROTQVLGL--TQTCETLKQ 718
Db 498 CFYTYGNYLPDHDIDISYTVTVDSGIIANRRAMFVDMSEITIKRRLAVSTQCFDLRAY 557
Qy 719 LNCIEDPVSPIVLRNLFNFSLVGTPLSAFGN-----LRPVLAEADAQRLFTALPFFEK 769
Db 558 VGNISIEDKLTPIKVTLQYDL-----NNDESRLQPHLELPIDNATMSTQTKQVSION 609
Qy 770 NCNDNHCODDLISITFSFMSLDCLVVGGPREFNVTVVRNDGDSYRTQVTFPPFLDLSY 839
Db 610 NCVN-NICIPDLDTVT-PNLPNIVIGQTELTDVSLNRRGEDAFQSSLSVYPLQLQF 667
Qy 830 RKVSTLQNRQSRQSWRLACESASSTVSGALKSTCSINHPIPEN-----SEVTNIT 883
Db 668 VRU-----ERKANMDFSVTCESD-----LRITCDTGNPMVGKNILEFGLTSLTFQVS 717
Qy 884 FOVDS-----KASLGNKLLKANVTSENNMPTNKTEFQLELPVKVAVMVVTSHGVS TK 938
Db 718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDCITLKLSSASYPEIVMYSTQED 775
Qy 939 YL-----NFTASENTSRVMQHOYQV-----SNLQORSLPIS-----LVFLVPV 976
Db 776 YVVPFPFPAKNASEADICMEVM-HLYEVRNTGSSNAGEVSLNIQMPQKNEDGEYLFYLLGI 834
Qy 977 RLNQTVINDRPO-----VTFSENLSSTCHTKERLPSHSDFLAELRKA PVV 1021
Db 835 MTEEGVTCQLTQOKANPEGVKLEPSTKAKLSNSETTQVSGRKRREPEVAEALQTDN--VI 892
Qy 1022 NCSTIACVQRIQCDDIPFFGQIEEFNAT-----LKGNLSFDWYIKTSHNHLIIVSTAEILF 1075
Db 893 YCASDSCVLINCTI-----DEINASKSVKVRILGRF--W-----ERTF 928
Qy 1076 NDSVFTLLPGQGAFAVRSQETKVE--PFEVNP-----LP----- 1108
Db 929 QKAVSELTPVQVQATIASSASAAVKTIPIYNIPLPRDFSSTKASTLVTTTELVPVPTPIAW 988
Qy 1109 -LIVGSSVGLLLALITAAALYKLGFFKR-----QYKDMMS--EGGPP 1148
Db 989 WIIIVSVLGGIILLLIILGLWKCGFFERKKPGEKEEYAPVASADKGGPP 1038
```

Search completed: November 25, 2003, 14:21:50
Job time : 21.1684 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.30328 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481B-1
Perfect score: 5956
Sequence: 1 MALRVLLTALTLCGFNLD.....FKRQYKDMSEGGPGGAEPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5940.5	99.7	1152	ITAM_HUMAN	P11215 homo sapien
2	4545	76.3	1153	ITAM_MOUSE	P05555 mus musculus
3	3503	58.8	1163	ITAX_HUMAN	P20702 homo sapien
4	3455	58.0	1162	ITAD_HUMAN	Q13349 homo sapien
5	1563	26.2	1170	ITAD_HUMAN	P20701 homo sapien
6	1549	26.0	1163	ITAM_MOUSE	P24063 mus musculus
7	1165.5	19.6	1167	ITAE_MOUSE	Q60677 mus musculus
8	1163	19.5	1179	ITAE_HUMAN	P38570 homo sapien
9	1108.5	18.6	1151	ITAH_HUMAN	P56199 homo sapien
10	1102.5	18.5	1189	ITAH_HUMAN	Q9UKX5 homo sapien
11	1099	18.5	1170	ITAH_BOVIN	P53710 bos taurus
12	1089	18.3	1178	ITAH_MOUSE	Q62469 mus musculus
13	1087.5	18.3	1181	ITAH_HUMAN	P17301 homo sapien
14	1085.5	18.2	1167	ITAH_HUMAN	O75578 homo sapien
15	1082	18.2	1180	ITAH_RAT	P18614 rattus norv
16	667	11.2	1039	ITAH_MOUSE	Q00651 mus musculus
17	638	10.7	1038	ITAH_HUMAN	P13612 homo sapien
18	630	10.6	1035	ITAH_HUMAN	Q13797 homo sapien
19	600	10.1	1032	ITAH_XENLA	Q91687 xenopus lae
20	571.5	9.6	1066	ITAH_CRISP	P17852 cricetidae
21	567.5	9.5	1053	ITAH_MOUSE	Q62470 mus musculus
22	555.5	9.3	1053	ITAH_MOUSE	P11688 mus musculus
23	550	9.2	1034	ITAV_CHICK	P26008 gallus gall
24	546.5	9.2	1050	ITAH_XENLA	Q08274 xenopus lae
25	540	9.1	1130	ITAH_HUMAN	P23229 homo sapien
26	535	9.0	1044	ITAV_MOUSE	P43406 mus musculus
27	532	8.9	1049	ITAH_HUMAN	P08648 homo sapien
28	531.5	8.9	1066	ITAH_HUMAN	P26006 homo sapien
29	530.5	8.9	1072	ITAH_CHICK	P26007 gallus gall
30	526	8.8	1048	ITAV_HUMAN	P06756 homo sapien
31	517	8.7	1044	ITAH_CHICK	P26009 gallus gall
32	512.5	8.6	1091	ITAH_MOUSE	Q61739 mus musculus
33	508	8.5	1396	ITAH_DROME	P12080 drosophila

RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol. chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2159(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
EX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

ALIGNMENTS

34	498	8.4	1179	1	ITAT_MOUSE	Q61738 mus musculus
35	493	8.3	1033	1	ITAB_MOUSE	Q9qum0 mus musculus
36	491.5	8.3	1146	1	ITAT_DROME	Q24247 drosophila
37	489.5	8.2	1039	1	ITAB_HUMAN	P08514 homo sapien
38	489	8.2	1025	1	ITAB_HUMAN	P51708 homo sapien
39	486	8.2	126	1	ITAM_CAYPO	P11578 cavia porce
40	474.5	8.0	1226	1	PAT2_CABEL	P34446 caenorhabdi
41	472	7.9	1181	1	ITAT_HUMAN	Q13683 homo sapien
42	469.5	7.9	1106	1	ITAT_RAT	Q63258 rattus norv
43	454.5	7.6	1139	1	INAT_CABEL	Q03600 caenorhabdi
44	424.5	7.1	1115	1	ITAT_DROME	O44386 drosophila
45	383	6.4	1000	1	ITAT_DROME	Q9wlm8 drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaut M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Paul H.L., Rosmarin A.G., Iken D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaut M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol:
conservation across species and homology to platelet IIb/IIIa.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
assessment.";
RL Structure 6:923-935(1998).
RN [12]
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Orvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
OF FIBRINOGEN GAMMA CHAIN.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
GRANULOCYTES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S52227; AAB24821.1; -
DR EMBL; S52152; AAB24821.1; JOINED.
DR EMBL; S52153; AAB24821.1; JOINED.
DR EMBL; S52154; AAB24821.1; JOINED.
DR EMBL; S52155; AAB24821.1; JOINED.
DR EMBL; S52157; AAB24821.1; JOINED.
DR EMBL; S52159; AAB24821.1; JOINED.
DR EMBL; S52161; AAB24821.1; JOINED.
DR EMBL; S52164; AAB24821.1; JOINED.
DR EMBL; S52165; AAB24821.1; JOINED.
DR EMBL; S52167; AAB24821.1; JOINED.
DR EMBL; S52169; AAB24821.1; JOINED.
DR EMBL; S52170; AAB24821.1; JOINED.
DR EMBL; S52173; AAB24821.1; JOINED.
DR EMBL; S52174; AAB24821.1; JOINED.
DR EMBL; S52180; AAB24821.1; JOINED.
DR EMBL; S52181; AAB24821.1; JOINED.
DR EMBL; S52184; AAB24821.1; JOINED.
DR EMBL; S52189; AAB24821.1; JOINED.
DR EMBL; S52191; AAB24821.1; JOINED.
DR EMBL; S52192; AAB24821.1; JOINED.
DR EMBL; S52203; AAB24821.1; JOINED.
DR EMBL; S52212; AAB24821.1; JOINED.
DR EMBL; S52213; AAB24821.1; JOINED.
DR EMBL; S52216; AAB24821.1; JOINED.
DR EMBL; S52219; AAB24821.1; JOINED.
DR EMBL; S52220; AAB24821.1; JOINED.
DR EMBL; S52221; AAB24821.1; JOINED.
DR EMBL; S52222; AAB24821.1; JOINED.
DR EMBL; S52226; AAB24821.1; JOINED.
DR EMBL; W76724; AAA58410.1; -
DR EMBL; M84477; AAA51960.1; -
DR PIR; A31108; RMHUIB.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWF; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.

Query Match		99.7%	Score 5940.5;	DB 1;	Length 1152;		
Best Local Similarity		99.9%	Pred. No. 0;				
Matches 1152; Conservative		0;	Mismatches	0;	Indels	1;	Gaps 1;
QY	1	MALRVLLLTALTLCHGFNLDENAMTFQENARGFGSVVLOGSRVVGAPQEIIVAAQR	60				
DB	1	MALRVLLLTALTLCHGFNLDENAMTFQENARGFGSVVLOGSRVVGAPQEIIVAAQR	60				
QY	61	GSLYQCDXSTGCEPIRLQVPEAVNMSLGLSLAATTSPPQLLAGCPTVHQTCSNTYVK	120				
DB	61	GSLYQCDXSTGCEPIRLQVPEAVNMSLGLSLAATTSPPQLLAGCPTVHQTCSNTYVK	120				
QY	121	GLCFLFGSNLRQOPKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL	180				
DB	121	GLCFLFGSNLRQOPKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL	180				
QY	181	KSKKTLFSLMOYSEFRTHFTKEFQNNPNPSLVKPIITQLGRTHATGIRKVVRELFN	240				
DB	181	KSKKTLFSLMOYSEFRTHFTKEFQNNPNPSLVKPIITQLGRTHATGIRKVVRELFN	240				
QY	241	ITNGARKNAFKILVITDGEKFDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQEL	300				
DB	241	ITNGARKNAFKILVITDGEKFDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQEL	300				
QY	301	NTIASKPPRDHVFQNNFEALKTIONQUREKIIFAEITGTGSSSFHEHMSQEGFSAIT	360				
DB	301	NTIASKPPRDHVFQNNFEALKTIONQUREKIIFAEITGTGSSSFHEHMSQEGFSAIT	360				
QY	361	SNGPLLSTVGSVDMAGGVFLYTSKSTFINNTRVDSMDNDAYLGAAIILNRVQSLV	420				
DB	361	SNGPLLSTVGSVDMAGGVFLYTSKSTFINNTRVDSMDNDAYLGAAIILNRVQSLV	420				
QY	421	LGAPRYQHILGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP	480				
DB	421	LGAPRYQHILGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP	480				
QY	481	HYEOTRGQVSVCPPLRGORARWQCDAYLGEQOQWGRFGAALTIVLGDVNGDKLTDVA	540				
DB	481	HYEOTRGQVSVCPPLRGORARWQCDAYLGEQOQWGRFGAALTIVLGDVNGDKLTDVA	539				
QY	541	ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLQYFGSLSGGQDLTMDGLV	600				
DB	540	ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLQYFGSLSGGQDLTMDGLV	599				
QY	601	DLTVAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVQK	660				
DB	600	DLTVAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVQK	659				
QY	661	STRDLREGQIOSVVTYDLALDSGRPHSAFNETKNSTRRTQVLGLTQTCETIKLQLP	720				
DB	660	STRDLREGQIOSVVTYDLALDSGRPHSAFNETKNSTRRTQVLGLTQTCETIKLQLP	719				
QY	721	NCIEDPVSPIVRLNFSVLGTPLSAFGNLRPVLAEQAQLFTALPFENKNCNDNICQDD	780				
DB	720	NCIEDPVSPIVRLNFSVLGTPLSAFGNLRPVLAEQAQLFTALPFENKNCNDNICQDD	779				
QY	781	LSITFSMSLDCLVGGPREFNVTVTRNDGSDSVRTQVTFEPFLDLSVRKVTSTLQNRS	840				
DB	780	LSITFSMSLDCLVGGPREFNVTVTRNDGSDSVRTQVTFEPFLDLSVRKVTSTLQNRS	839				
QY	841	QRSWLACSSASTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA	900				
DB	840	QRSWLACSSASTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA	899				
QY	901	NVTSENMPRTNKTETPQLPVLKYAVYVVTSHGVSTKYLNFTASNTSRVNHQVQSN	960				
DB	900	NVTSENMPRTNKTETPQLPVLKYAVYVVTSHGVSTKYLNFTASNTSRVNHQVQSN	959				
QY	961	LGQRSLPISLFLVPVRLNQTWIDRPQVTFSENLSSTCHTKERLPSPSHDFLAEARKAPV	1020				
DB	960	LGQRSLPISLFLVPVRLNQTWIDRPQVTFSENLSSTCHTKERLPSPSHDFLAEARKAPV	1019				

1021 VNCIAVCORIQCIPPFQIOEFNATLKGNSLSDWYIKTSHNHLIIIVSTAILFNDVSF 1080
1020 VNCIAVCORIQCIPPFQIOEFNATLKGNSLSDWYIKTSHNHLIIIVSTAILFNDVSF 1079
1081 TLLPQGAQFVSQSTETKVEPEVNPPLPIVGVSSVGGLLALLLITAAALYKLGFFFRQYKD 1140
1080 TLLPQGAQFVSQSTETKVEPEVNPPLPIVGVSSVGGLLALLLITAAALYKLGFFFRQYKD 1139
1141 MWSEGGPPEGAPQ 1153
1140 MWSEGGPPEGAPQ 1152

RESULT 2
ITAM MOUSE
ID ITAM MOUSE STANDARD; PRT: 1153 AA.
AC P05555;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).
GN ITGAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89312584; PubMed=3044779;
RA Pyeola R.;
RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";
RL EMBO J. 7:1371-1378 (1988).
RN [2]
RP SEQUENCE OF 11-45 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=86287312; PubMed=2942940;
RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;
RT "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648 (1986).
RN [3]
RP SEQUENCE OF 17-28.
RX MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542 (1985).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO AN IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

Qy 1080 FTLPGGQAFVRSQTKVPEFFVNPVLPVIVGSSVGLLLALITAAKYKLGFFKQYK 1139
 Db 1080 FALLPGQESYVRSQTKVPEVHNVPVIVGSSIGLVLLALITAGLYKLGFFKQYK 1139
 Qy 1140 DMNSEGGPPGAEPO 1153
 Db 1140 DMNEAAPQDAPPO 1153
 RESULT 3
 ITAX_HUMAN
 ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
 AC P20702;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
 DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
 GN ITGAX OR CD11C.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RN SEQUENCE FROM N.A.
 RP MEDLINE=88166645; PubMed=3327687;
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
 RT "cDNA cloning and complete primary structure of the alpha subunit of
 RT a leukocyte adhesion glycoprotein, p150,95.";
 RL EMO J. 6:4023-4028(1987).
 RN (2)
 RN SEQUENCE FROM N.A.
 RP MEDLINE=90153906; PubMed=2303426;
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RT "Genomic structure of an integrin alpha subunit, the leukocyte
 RT p150,95 molecule.";
 RL J. Biol. Chem. 265:2782-2788(1990).
 RN (3)
 RN ERRATUM.
 RP Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RL J. Biol. Chem. 265:12750-12751(1990).
 RN (4)
 RN SEQUENCE OF 20-43.
 RX MEDLINE=87167596; PubMed=3549901;
 RA Miller L.J., Wiebe M., Springer T.A.;
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1
 RT and p150,95 leukocyte adhesion proteins.";
 RL J. Immunol. 138:2381-2383(1987).
 CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
 CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
 CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
 CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
 CC ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
 CC GRANULOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VFMA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----

DR ENBL; M81695; AA59180.1; --
 DR ENBL; Y00093; CAA68283.1; --
 DR ENBL; M29165; --; NOT ANNOTATED CDS.
 DR ENBL; M29487; AA51620.1; ALT SEQ.
 DR ENBL; M29482; AA51620.1; JOINED.
 DR ENBL; M29483; AA51620.1; JOINED.
 DR ENBL; M29484; AA51620.1; JOINED.
 DR ENBL; M29485; AA51620.1; JOINED.
 DR ENBL; M29486; AA51620.1; JOINED.
 DR PIR; A36584; RWHUIC.
 DR PDB; 1N3Y; 18-FEB-03.
 DR Genew; HGNC:6152; ITGAX.
 DR MIM; 151510; --
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0004872; F:receptor activity; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWF_A; 1.
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 1163
 FT DOMAIN 20 1107
 FT TRANSMEM 1108 1128
 FT DOMAIN 1129 1163
 FT REPEAT 34 87
 FT REPEAT ? ?
 FT DOMAIN 165 351
 FT REPEAT ? ?
 FT REPEAT 402 453
 FT REPEAT 455 517
 FT REPEAT 518 576
 FT REPEAT 581 633
 FT CA_BIND 466 474
 FT CA_BIND 530 538
 FT CA_BIND 593 601
 FT SITE 1131 1135
 FT DISULFID 69 76
 FT DISULFID 108 126
 FT DISULFID 655 712
 FT DISULFID 771 777
 FT DISULFID 848 863
 FT DISULFID 998 1022
 FT DISULFID 1027 1032
 FT CARBOHYD 61 61
 FT CARBOHYD 89 89
 FT CARBOHYD 392 392
 FT CARBOHYD 697 697
 FT CARBOHYD 735 735
 FT CARBOHYD 899 899
 FT CARBOHYD 939 939
 FT CARBOHYD 1050 1050
 FT CONFLICT 490 490
 FT CONFLICT 756 756
 SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19C3F62A473 CRC64;
 Query Match 58.8%; Score 3503; DB 1; Length 1163;
 Best Local Similarity 61.3%; Pred. No. 2,7e-229;
 Matches 700; Conservative 138; Mismatches 237; Indels 6; Gaps 4;
 Qy 5 VLLTALTLCGHFNLDTENAMTFQENAGFGQSVVLQGGSRVVVGAPOEIVAAHQSGLS 64
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 8 LLLFTALATSLGPNLDTELTAPRVDSAGFGDSVVQVANSVVVGAPOKITAANQTGGLY 67

Db 901 ENNKASSKATFOLELKVYAVVTMISRQESTKYPNFATSDKKMKEABHRYVNNLSQ 960
 Qy 964 RSPISLVLPVRLNQTIVDRPQVTFSENLSCTHKERLSDHSDFLAELKAPVNC 1023
 Db 961 RLDAISFNFWPVLNGLVAVVMEAPQSLL--PCVSRKPPQSDFLTQISRPMLDC 1018
 Qy 1024 SIACVORIQCIDIPFFQEFNATLKGNSLDFWIKTSHNHLIVSTAEILFNDVFTLL 1083
 Db 1019 SIADCLQFRCDVFSFQEBLDFLKGNSLFGWVRETLQKVLVSVAEITFTSVYSQL 1078
 Qy 1084 PQGAFFVRSQETKTEFFEPVNPPLPLVSGVGLLALLITAAALYKLGFFKQYKDMMS 1143
 Db 1079 PQGAFFVRSQETKTEFFEPVNPPLPLVSGVGLLALLITAAALYKLGFFKQYKDMMS 1143
 Qy 1144 E 1144
 Db 1139 D 1139

RESULT 5
 ITAL_HUMAN STANDARD; PRT; 1170 AA.
 AC P20701; O43746;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
 DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)
 DE (CD11a).
 GN ITGAL OR CD11A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
 RX MEDLINE=89139587; PubMed=2537322;
 RA Larson R.S., Corbi A.L., Berman L., Springer T.;
 RT "Primary structure of the leukocyte function-associated molecule-1
 RT alpha subunit: an integrin with an embedded domain defining a protein
 RT superfamily.";
 RL J. Cell Biol. 108:703-712(1989).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=99425270; PubMed=10493829;
 RA Loftus B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R.,
 RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
 RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
 RT "Genome duplications and other features in 12 Mb of DNA sequence from
 RT human chromosome 16p and 16q.";
 RL Genomics 60:295-308(1999).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
 RX MEDLINE=96036067; PubMed=7479767;
 RA Qu A., Leahy D.J.;
 RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha
 RT L beta 2) integrin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
 RX MEDLINE=96398682; PubMed=8805579;
 RA Qu A., Leahy D.J.;
 RT "The role of the divalent cation in the structure of the I domain
 RT from the CD11a/CD18 integrin.";
 RL Structure 4:931-942(1996).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
 RX MEDLINE=99425288; PubMed=10493852;
 RA Kallen J., Weizenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
 RA Cottens S., Weitz-Schmidt G., Hommel U.;
 RT "Structural basis for LFA-1 inhibition upon lovastatin binding to the
 RT CD11a I-domain.";

RL J. Mol. Biol. 292:1-9(1999).
 CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
 CC ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
 CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
 CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
 CC AND MONOCYTES.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
 CC ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P20701-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P20701-2; Sequence=VSP_002738;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y00796; CAA68747.1; --
 DR EMBL; AC002310; AAC31672.1; --
 DR PIR; S03308; S03308.
 DR PDB; 1LFA; 29-JAN-96.
 DR PDB; 1ZON; 07-DEC-96.
 DR PDB; 1ZOO; 07-DEC-96.
 DR PDB; 1ZOP; 07-DEC-96.
 DR PDB; 1CQP; 07-AUG-00.
 DR PDB; 1DQJ; 03-FEB-00.
 DR PDB; 1MJN; 28-JAN-03.
 DR PDB; 1MOB; 14-JAN-03.
 DR PDB; 1MQ9; 14-JAN-03.
 DR PDB; 1MQA; 14-JAN-03.
 DR Genem; HGNC:6148; ITGAL.
 DR MIM; 153370; --
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0006928; P:cell motility; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; 3D-structure; Magnesium; Calcium; Repeat;
 KW Alternative splicing.
 FT SIGNAL 1 25
 FT CHAIN 1 25
 FT DOMAIN 26 1170 INTEGRIN ALPHA-L.
 FT TRANSMEM 1089 1112 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1113 1170 POTENTIAL.
 FT REPEAT 42 91 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 92 149 FG-GAP 1.
 FT DOMAIN 170 349 FG-GAP 2.
 FT REPEAT ? ? FG-GAP 3.

RESULT 6

ITAL MOUSE STANDARD; PRT: 1163 AA.

AC P24063;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).

GN ITGAL OR LFA-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91268576; PubMed=2051027;

RA Kaufmann Y., Teeng E., Springer T.A.;

RT "Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit and its expression in COS cells.";

RL J. Immunol. 147:369-374(1991).

RN [2]

RP SEQUENCE OF 24-42.

RX MEDLINE=85188276; PubMed=3887182;

RA Springer T.A., Teplow D.B., Dreyer W.J.;

RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";

RL Nature 314:540-542(1985).

CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2, ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED LEUKOCYTES RECRUITMENT.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L ASSOCIATES WITH BETA-2.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: LEUKOCYTES.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A WFPA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: Contains 1 WFPA domain.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M60778; AAA39426.1;

DR PIR; I56126; I56126.

DR HSPD; P20701; 1LFA.

DR MGD; MGI:96606; Itgal.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; WVF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin_A; 1.

DR Pfam; PF00092; Wva; 1.

DR PRINTS; PR01185; INTEGRINA.

DR SMART; SM00453; WVFADOMAIN.

DR SMART; SM00191; Int_alpha; 5.

DR SMART; SM00327; WVA; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS0234; WVFA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane; Signal; Magnesium; Calcium;

KW Repeat. 1 23
 FT SIGNAL 24 1163
 FT CHAIN 24 1084
 FT DOMAIN 1085 1108
 FT TRANSMEM 1109 1163
 FT DOMAIN 39 88
 FT REPEAT ?
 FT REPEAT ?
 FT DOMAIN 148 334
 FT REPEAT ?
 FT REPEAT 399 454
 FT REPEAT 455 514
 FT REPEAT 516 573
 FT REPEAT 576 628
 FT CA_BIND 466 474
 FT CA_BIND 528 536
 FT CA_BIND 588 596
 FT SITE 1111 1115
 FT DISULFID 70 77
 FT DISULFID 108 126
 FT DISULFID 147 199
 FT DISULFID 651 705
 FT DISULFID 767 773
 FT DISULFID 840 856
 FT DISULFID 993 1009
 FT DISULFID 1017 1048
 FT CARBOHYD 86 86
 FT CARBOHYD 185 185
 FT CARBOHYD 270 270
 FT CARBOHYD 444 444
 FT CARBOHYD 668 668
 FT CARBOHYD 696 696
 FT CARBOHYD 724 724
 FT CARBOHYD 728 728
 FT CARBOHYD 776 776
 FT CARBOHYD 857 857
 FT CARBOHYD 880 880
 FT CARBOHYD 890 890
 FT CARBOHYD 899 899
 FT CARBOHYD 927 927
 FT CARBOHYD 1056 1056
 SQ SEQUENCE 1163 AA; 128343 MW; A7A307848988232F CRC64;

Query Match 26.0%; Score 1549; DB 1; Length 1163;
 Best Local Similarity 34.2%; Pred. No. 8 4e-97;
 Matches 409; Conservative 216; Mismatches 464; Indels 106; Gaps 38;

QY 1 MALRV-----LLLTALT-----CHGFNLDTENAMTFQENA-RGFGQSVVQLQSGRVVVGAPQ 52
 DB 1 MSFRIAGPRLLLGLQLPAKWSYNLDRPTQSFQAQGRHFGYQLQIEDG-VVVGAPG 59
 QY 53 EIVANQSGSLYQCDYSTGSCPEIRLOVPVEAVNWSLGLSLAATTSPPOLLACGPTVHOT 112
 DB 60 E---GDNTGGLYHCRTSSEFCQPVSLH-GSNHTSKYLGMTLATDAAKGSLLACDPLGSR 115
 QY 113 CSENTYVVGKLCFLFGSNLRQQPKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEF 172
 DB 116 CDQNTYLSGLCYLFPQSLGPMQNPAYQCKGKGVLDVFLFDGSQLDKDKFEKILEF 175
 QY 173 VSTVMEQLKSKTFLSLMQYSEEFRIHFTKFE-QNNPNRSLVKPIETOLLGTHHTATGI 231
 DB 176 MKDVMKLSNTSYQFAAVQFSTDCRTEFTFLDYVKQKNKPNVDVLLGSGVQPMFLLTNTFRAI 235
 QY 232 RKVRELFINITNGARKNAFKILVITDGEKFGDPLGYEDVIPADREG-----VIRYV 284
 DB 236 NYVAHVFKESGARDPAATKVLVIITDG-----EASDKGNISAAHDITRYI 281
 QY 285 IGVGDAFRSEKSRQELNTIASPPRDHVQVNNFEALKTIONOLREKIFAIEGTQTGSSS 344
 DB 282 IGKHKHVSQKQKTHIFASEVEFEVKILDTFEKLDKDLQRLQRIAYIEGTNRQDLT 341
 QY 345 SFEHEMSQEGFSAITSNGPFLSTVGSDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAY 403

Db 342 SFNMLSSSGISADLSKSHAVGVAGAKDWAGGLDLREDLQAGTFVQGBLTSVDRGGY 401
Qy 404 LGYAAA-IILNRVQSLVGLAPRYQHIGLVAMFR--QNTGMWESNANYKGTQIGAYFGAS 460
Db 402 LGYTVAMTSSRSLPAAAPRYOHVQVLLFOAPEAGGRWNTQKTEGTQIGSYFGGE 461
Qy 461 LCSVDVDSNGSTDLVLIAGPHYETGTRGGQVSVCLPRGQARWOCDAVLVGEQOPWGR 520
Db 462 LCSVDLDQGEAEELLIGAPLFFGQGRGRVFTY---QRROSLFEMVSELQDGPYPLGR 518
Qy 521 FGAALTVLGDVNGDKLTVAIGAPCEEDNRGAVLFGHTSGSISPSHSQRIAGSKLSPR 580
Db 519 FGAALTALTDINGDLTLDVAVGAPLEE--QGAVYIFNKGPG-GLSPQSPQRIQGAQVFP 575
Qy 581 LOYFQCSLGGQDLTMDGLVLTGVAQGHVLLRSQPLRVKAINMEFNPVARNVFCN 640
Db 576 IRWFGRSHGVKDLGDLADLVVCAEGRVVVLSRPVVDVVTLSFSPPEEIPVHEVCS 635
Qy 641 DQVVGKENG-EVRVCLHVQKSTRDLREGQIQSVVTDLALDSORPHSRVAFNETKNT 699
Db 636 YSAREEQHGVKLCACFRKPLTPQ--FQGRLLANLSYTLQDGHMRSGLFPDGSHEL 693
Qy 700 RQQTQVLGTQTCETLKLQPLNCIEDPVSPIVLRNFSIV---GTPLSAFGN-LRPVLA 755
Db 694 SNTSITP-DKSCDLDFHFFPICQDLISPINVSFNLSLEEGRTPDQKGRAMQPILRP 752
Qy 756 DAQRFTALFPPEKNCNDNICQDLSITFSFMSLDCLVGGP-----REFNVTVTVRN 809
Db 753 SIHTV-TKEIPPEKNCGEDKCEANLTLSPPARS-----GPLRLSSASLAVEMTILSN 804
Qy 810 DGEDSYRTQVTFPFLDLSYRKVSTLQNRORSWRLACESASSTEVSGAL-KSTSCSIN 868
Db 805 SGEDAYVWRDLDFPRGSLFRKVEMLQ---PHSRMPVSCSEL--TEGSSLLTKLTKCNVS 859
Qy 869 HPFENSEVTNITFDVDSKASLGNKLLKANVTSEN-NMPTNKTEFQLELPVYKXVY 927
Db 860 SPIFKAGEVSVQVFMFTLLNSWEDFVELNLTGTVHCENENSSQLQEDNSAATHIPVLPVN 919
Qy 928 MVVTSHTGVTYKLTNTASENTSRVMOHQVQVNLGQSLPISLVLVPLVRLNQTIVMDRP 987
Db 920 ILTKQENSTLYISTPKGPKTQOVHVYQV-----RQPSAYDNMPT-LEALVGVPRP 973
Qy 988 Q-----VTSENLSS-----TCHTKB-RLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIPFF 1038
Db 974 HSEDLTITVWSQTDPLVTCHESDLKRPSE---AEQPCPLGV-----QFRCPVIF- 1021
Qy 1039 GQOEFENATLKNLSFDMVYIKTSHNHLIVSTAELFNDLSVFTLLPGQAFVRSQTETKV 1098
Db 1022 --RWEILIQVGTVELSKIEIKAS-STLSLCSLSVSFNSSKHFHLYGSKA-SEAQVLVKV 1077
Qy 1099 EPFEVNPPLVIGSVSGVGLLLALITAAALYKLGFFKQYKDMW-SEGPPGCAEP 1152
Db 1078 DLHHEKMLHVYVLSGIGGLVLLFLFLALYKVGFFKRLKEMKMEADGVNPGSP 1132

RESULT 7
ITAE_MOUSE STANDARD; PRT; 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN ITGAE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKR;
RX MEDLINE=95187992; PubMed=7882170;
RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,

RA Kilshaw P.J., Weis J.H.;
RT "Murine M290 integrin expression modulated by mast cell activation.";
RL Immunity 1:393-403(1994).
CC !- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC !- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC !- SUBCELLULAR LOCATION: Type I membrane protein.
CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC !- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC !- SIMILARITY: Contains 1 VMFA domain.
CC !- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U12236; AAC52142.1; --
CC HSP: P11215; 1A8X.
CC MGD: MGI:1298377; Itgae.
CC InterPro: IPR000413; Integrin_alpha.
CC InterPro: IPR002035; VWF_A.
CC Pfam: PF01839; FG-GAP; 3.
CC Pfam: PF00357; integrin_A; 1.
CC Pfam: PF00092; vwa; 1.
CC PRINTS: PR01185; INTEGRINA.
CC SMART: SM00191; Int_alpha; 3.
CC SMART: SM00453; VWFADOMAIN.
CC SMART: SM00327; VWA; 1.
CC PROSITE: PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE: PS02334; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Magnesium;
CC Calcium.
CC SIGNAL 1 19 BY SIMILARITY.
CC CHAIN 20 1167 INTEGRIN ALPHA-E.
CC CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.
CC CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.
CC DOMAIN 20 1114 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1115 1137 POTENTIAL.
CC DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).
CC REPEAT ? ? FG-GAP 1.
CC REPEAT ? ? FG-GAP 2.
CC REPEAT ? ? X-DOMAIN (EXTRA DOMAIN).
CC DOMAIN 149 192 VWFA.
CC DOMAIN 193 384 FG-GAP 3.
CC REPEAT ? ? FG-GAP 4.
CC REPEAT 449 501 FG-GAP 5.
CC REPEAT 503 564 FG-GAP 6.
CC REPEAT 566 631 FG-GAP 7.
CC REPEAT 634 686 POTENTIAL.
CC CA_BIND 514 522 POTENTIAL.
CC CA_BIND 578 586 POTENTIAL.
CC CA_BIND 646 654 POTENTIAL.
CC DOMAIN 185 191 GLU-RICH (ACIDIC).
CC SITE 1140 1144 GFFKR MOTIF.
CC DISULFID 72 83 BY SIMILARITY.
CC DISULFID 130 164 BY SIMILARITY.
CC DISULFID 698 754 BY SIMILARITY.
CC DISULFID 814 820 BY SIMILARITY.
CC DISULFID 884 898 BY SIMILARITY.
CC DISULFID 998 1023 BY SIMILARITY.

FT DISULFID 1031 1047 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 829 829 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 846 846 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFD CRC64;
Query Match 19.68; Score 1165.5; DB 1; Length 1167;
Best Local Similarity 29.18; Pred. No. 8.66-71;
Matches 362; Conservative 213; Mismatches 465; Indels 205; Gaps 43;
QY 5 VLLTALTLCGFLNLTENA--MTFQENARGFGOSVVOLOQSRVVGAPQEIIVAAANQSGS 62
DB 8 LLMASLKPGAFNMDVDWAWTALQCAPAVLSLLHLDPSN-----NOTCLLVARRSS 62
QY 63 -----LYQCDYSTGSCPEIRLQVPEAVNMNLSGLSLAATT--SPQLLAC-GPTVHQT 113
DB 63 NRNTAALYRCAISI-SPEIACQ-PVEHICMPKGRYQGVTLVGNHNGVLCIQVOARKPR 120
QY 114 SENTYVGLCLFSGNLRQOKPEALRG-----C----- 144
DB 121 SLNSELTCACSLTPNLDLQQAAYFSDLEGFLDPCAHAHVDGSDYCRSKSGSTGEBKKSARR 180
QY 145 -----POEDSDIAFLIDSGSIIPHDFRRMKEFVSTVMEQL--KSKTLFLSLMOYS 193
DB 181 RTVEEEDDEEDGETEIAIVLDGSGSIGSPDFQKAKNISTMRNFYEKCFECFNALVQYG 240
QY 194 EEFRIHTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAPKIL 253
DB 241 AVIQTEPDQSRDINASLAKVQSIQVQKVTASAMQHVLDFNIFPSRGRKALKVM 300
QY 254 VVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSGQELNTTASKEPRHVF 313
DB 301 VVLTGDGDFGDLPLTTVINSFPMQGVVFAIGVGRFKNNTYRELKLIASDPKEAHTF 360
QY 314 QVNNFEALKTIONLRKIFALEGTQTGSSSSFEHMSQEGFSAITSNGP-LLSTVGSY 372
DB 361 KVTNYSALDGLSLKQRIQVHMEGT---VGDALQQLAQTGFSQIILDKGQVLLGTVGAF 417
QY 373 DWAGGVFLY-TSKBKSTFINMT-RVDS-DMNDAYLVGAALILNRVQSLVGLGAPRYOHI 429
DB 418 NWSGALLYSTONGRCFLNQTAKEDESRVQSYVGLSVLAVLHKAHGISYVAGAPRHKL 477
QY 430 GLVAFRONTGMWESNA---NVKGTQIAYFGASLCSVDVDSNGSTDLVLICAPHYEQT 486
DB 478 GAVFELRKEDR--BEDAFVRRIEGEQMGSGFVSLCPVDIDMDGTDFLLVAAPFFYHIG 535
QY 487 RGGGVSVCLPRGQARQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAP-- 544
DB 536 BEGRVYVYQVPE-QDASFSLAHTLSGHPGLTNSRFGFAMAAVGDINQDKFTDVAIGAPLE 594
QY 545 ---GEEDNRGAVLHFGTSGSIGSPHSQRIAGSKLSPRLQYFGQSLSGQDLTMGLV 600
DB 595 GFGAGDGASGVVIYNGHSG-GLYDPSQIRASSVASGLHYFGMSVSGGLDFNGDLA 653
QY 601 DLTGVAQGHVLLLRSPQVLRVKAINEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQK 660
DB 654 DITVGSRDSAVVLRSPVVDLTVSMTFP-----DALPMVFYFGKM--DVNLCFEVDS 703
QY 661 S---TRDLRREGIOQSVTYDLALDSGRPHSAFVNETKNSRTRQTVLGLTQTC----- 712

DB 704 SVVASEPGLUREMFLNFTVDV-----TKQRQLQCESSGSCQCLRKWN 748
QY 713 -----ETLKLQLPNCIEDPVPSPVILVRLNFSVLGTPLSAFGLNR-----PVLAED 756
DB 749 GSGFLCEHFWLISTBEL-----CEEDCFSNITIKVTE-----FQTSGGRRDYPNPTL--D 797
QY 757 AORLFTALF--PFENKCGNDNICQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDS 814
DB 798 HYKEPSAIFQLPYEKDCKNKFVCAIEIQLTTN-ISOQLVVGVTKEVTMNSILTSNGEDS 856
QY 815 YRTQVTFPPDLVSRYKVTSTLQNRQSRWRACSSASSTEVSGALKSTSCSINHPIPPE 874
DB 857 YMTNMAALNYPRLQFKKI-----QKPVSPDVCQDCKPV---ASVLVNMCKIGHPII-K 906
QY 875 NSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTETFOLELPVKYAVYVMTVSHG 934
DB 907 RSVNVSVTWQLEESVFPNRTADITVTISNKSLSARETR---SLQFRAHIAVLNR-- 961
QY 935 VSTKYLNTASNTSRVMQHQYQVSNLQGRSLPISLVFLVLPVRLNQTIVWRPQVTFSEN 994
DB 962 PSVMYWN--TSQSPDHKEFFFNHVGENLFGAVFQLOICVPKLODF-----QIVRVKN 1013
QY 995 LGST-----CHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIPFFGIOEFPNATLK 1049
DB 1014 LTKTDHTECTQSEPCAGSDPVQHVKEHWSVVCAL-----TSNK 1053
QY 1050 GNLSPPWYIKTSHNHLIVSTA-----EILFNDVSFTLLPGCAFVRVSRQTEKVPFF-- 1101
DB 1054 ENVTVAEEISVGHKTOLRDVSELPIGSEISFNKSLYEGLNAB-----NHRTKITVIFL 1107
QY 1102 --EVPNPLPLIVGSSVGGLLALLALITAAALYKLGFFKRYQKDMSE 1144
DB 1108 KEETRSLPLITGSSIGLLLVVILAILFKCGFFKRYQKDMSE 1152
RESULT 8
ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and lymphocytes;
RX MEDLINE=9416496; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;
RA "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G., Shoclerauk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;
RA "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";

Genome. Res. 10:165-173 (2000).

[4]

RL MUTAGENESIS OF ASP-109 AND PHE-316.

RN MEDLINE=20400502; PubMed=10837471.

RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,

RA Brenner M.B., and beta chains in ligand recognition by beta 7

RT "The role of alpha and beta chains in ligand recognition by beta 7

RT integrins.";

RL J. Biol. Chem. 275:25652-25664 (2000).

CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT

CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL

CC CELL MONOLAYERS.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA

CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A

CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN

CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL

CC EPITHELIAL CELLS.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;

CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L55851; AAB59359.2; -

DR EMBL; AF168787; AAF34107.1; -

DR PIR; A53213; A53213.

DR HSSP; P11215; IABX.

DR Genew; HGNC:6147; ITGAE.

DR MIM; 604682; -

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; integrin_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 3.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PSS0234; VWFA; 1.

DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

DR Signal; Repeat; Polymorphism; Magnesium; Calcium.

FT SIGNAL 1 18

FT CHAIN 19 1179 INTEGRIN ALPHA-E

FT CHAIN 19 1177 INTEGRIN ALPHA-E LIGHT CHAIN.

FT CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.

FT DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1125 1147 POTENTIAL.

FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 181 198 GLU-RICH (ACIDIC).

FT REPEAT ? ?

FT REPEAT ? ? FG-GAP 1.

FT DOMAIN ? ? FG-GAP 2.

FT DOMAIN 145 199 X-DOMAIN (EXTRA DOMAIN).

FT DOMAIN 200 391 VWFA.

FT REPEAT 401 456 FG-GAP 3.

FT REPEAT 457 506 FG-GAP 4.

FT REPEAT 510 571 FG-GAP 5.

FT REPEAT 573 638 FG-GAP 6.

FT REPEAT 641 693 FG-GAP 7.

FT CA_BIND 522 530 POTENTIAL.

FT CA_BIND 586 594 POTENTIAL.

FT CA_BIND 654 662 POTENTIAL.

FT SITE 1150 1154 GPFKR MOTIF.

FT DISULFID 70 79 BY SIMILARITY.

FT DISULFID 126 159 BY SIMILARITY.

FT DISULFID 706 762 BY SIMILARITY.

FT DISULFID 823 829 BY SIMILARITY.

FT DISULFID 893 907 BY SIMILARITY.

FT DISULFID 1008 1033 BY SIMILARITY.

FT DISULFID 1041 1057 BY SIMILARITY.

FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 954 954 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARIANT 360 360 D -> E.

FT VARIANT 1041 1041 /FTID=VAR_008884.

FT VARIANT 208 208 /FTID=VAR_008885.

FT MUTAGEN 316 316 D->A: LOSS OF E-CADHERIN BINDING.

FT MUTAGEN 477 477 F->A: LOSS OF E-CADHERIN BINDING.

FT CONFLICT 482 482 V -> I (IN REF. 3).

FT CONFLICT 482 482 Q -> R (IN REF. 3).

FT CONFLICT 950 950 R -> W (IN REF. 3).

FT CONFLICT 1019 1019 A -> V (IN REF. 3).

SQ SEQUENCE 1179 AA; 130088 MW; E558902EDF9D95E1 CRC64;

Query Match 19.5%; Score 1163; DB 1; Length 1179;

Best Local Similarity 28.8%; Pred. No. 1.3e-70;

Matches 356; Conservative 221; Mismatches 481; Indels 180; Gaps 41;

QY 5 VLLLTALTLCGFLNLTENAMTFQENARGFGQSVVQLQG-----SRVVGVAQEIWAQ 59

DB 7 LLCIASLALLAAFNVDVAPWLTGKGFVLSLLHQDPSTNQWLLVTSR---TKRT 63

QY 60 RGSLYQCYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPPQLACGPTVH 110

DB 64 PGLHRCSLVQDEILCHPVEHVPKGRHGRVTVVRSHHGLICI-----QVLVRRP--H 116

QY 111 QTCSENTYVKGCLFLGSLNRQPOQ-----135

DB 117 SLSSSLT--GTCSLGPDLPQQAQNFLENLDPDARVDTGDCYSNKEGGEDDVNT 173

QY 136 -KFPEALRGCPQED-----SDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL- 180

DB 174 ARQRALEKEEEDKEEEDDEEEDEEAEAGTEAIIIDGSGSIDPDPQRAKDFISNMNRNY 233

QY 181 -KSKXTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELF 239

DB 234 EKCPFCNFAVQYGGVIOTEFDLRDSQDVMSLARVQMITQVGSVTKTASAMQHLDSIF 293

QY 240 NITNGARKNAFKILVITDGEKFGDPLGYEDVIPLEADREGVIRYVIGVDADFSEKSR 299

DB 294 TSSHGSRRKASKVMVVLTDGGIFEDPLNLTTVINSPKMGQVVERFAIGVEEFKSAITARE 353

QY 300 LNTIAKSPRDHVFQVNNFNAELKTIQNLREKIFAIEGTQTCSSSSSFEHMEQEGFSAAI 359

DB 354 LNLIASDPDETHAFKVTNYMALDGLLSKLRNIIISMEGT---VGDALYQLAQIGFSAOI 410

QY 360 TSNGP-LLSTVGSYDMAGGVFLY-TSEKSTFINNTRVDSMDNDA-----YLGYAAAIILR 413

DB 411 LDERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTAAAAADAAEAQSYLGAVAVLHK 470

QY 414 NRQSLVLGAPRYQHIGLVAMPR-ONTQGWESNANV-KGTQIGAVFGASLCSVDVDSNGS 471

DB 471 TCSLSYVAGAPQYKHG--AVPELOKQEGREASFLPVLEGEQMGSYFGSELCPVDIMDGS 528


```
Qy 472 TDLVLIGAPHYEQTRGGQVSCPLPRGQRARWOCDAVLXGEOQPMGRFGAALTILGVDV 531
Db 529 TDFLLVAAPFHHVHGEGRVYVYRLSE-QDGSFSLAILSGHFGFTNARGFAMAWGDL 587
Qy 532 NGDKLTDAVAGAP---GEEDNR--GAVYLFHGTSGSGISFHSQRTAGSKLSPRLQYFG 585
Db 588 SQDKLTDAVAGAPLEGFGADGAGSGFYIYNG-HWDGLSASPSQIRASTAVAPGLQYFG 646
Qy 586 QSLSGGQDLTWDGLVDLTGAGQGHVLLLRQPVLRVKAIEMFNPVARNVFCNDQOVK 645
Db 647 MSMAGGFDISGDGLADITVGLGOAVVFRSRRVRLKVSMAFTPSALP-----IG 696
Qy 646 GKEAGEVRVCLVHVKSTRDLREGQIOQSVVTDLDLSDGRPHSRAVFNETKNSRRQTQV 705
Db 697 FNGVNVRLCFEI--SSYTTASESGUREALLNFTLDVVGKRRRLQCSVRSCLGCLREW 755
Qy 706 LGLTQTCETLKLQLPN---CIEDPVPSPVLRNFSVLGTPLSAFGNLRPVLBAEDARLF 761
Db 756 SSGSQLCEDL--LLMPTTEGELCEEDCFNSVKSQYL-QTPEGQTDHPQPILDRYTEPFA 813
Qy 762 TALPFEKNCNDNICDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSVRTQVTF 821
Db 814 IFQLPYEKACKNKLFCVAELQLA-TTVSQOELVVLTKELTLNLTNSGEDSVTSMAL 872
Qy 822 FFLDLSYRKVSTLQNGRSQSRWRLACESASTEVSGALKSTSCSINHPIPPENSEVTFN 881
Db 873 NYPRNLQ-----LKMCKPSPNIOCDPOPV---ASVLMNCRIHGPVL-KESSAHVS 922
Qy 882 ITFDVDSKASLGNKLLKANVTSENN-----MPTNKTEFO---LELPKVAVMVVTSHG 934
Db 923 VVMOLEENAFNRTADIITVTNNGERSLANETHLQFRHGFAVLKSKPSIMVYNTGQG 982
Qy 935 VS--TKYLNFTASENTSRVMOHOVQVSNLQORSPLISLVFLVPVRLQTVTWDRPQVTF 992
Db 983 LSHKKEFLFHVHGN---LFGAEYQ-----LQICVPTKRLGLQVAAVKKLTRT 1027
Qy 993 ENLSSTCHTKERLPSHSDFLAELKAPVNVGSIACVQRIQCDIPFGIOEERFATLKGNL 1052
Db 1028 QASVTCTWSQERACAYSS-VQVHEWHSVSCVIA-----SDKENVTVAEI 1072
Qy 1053 SFDWYIKTSHNHLIVST-----AEILFNSVFTLLPGQAFVRSOTETKVEPPEVPNP 1106
Db 1073 SWD-----HSEELLKDVTELQILGEISFNKSLVEGLNAENH--RTKITVYFLDKERYHS 1124
Qy 1107 LPLIVGSSVGGILLALITALYKLGKFFKQYKQMMSE 1144
Db 1125 LPIIIGSVGGILLVILVILFKCGFFKRYQQLNLE 1162
```

RESULT 9

```
ITAL_HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.;
RT "Expression of native and truncated forms of the human integrin alpha
RT 1 subunit";
RL J. Biol. Chem. 268:2989-2996 (1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
CC E-R IN COLLAGEN.
```

```
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".
DR PIR; A45226; A45226.
DR PDB; 1QC5; 17-MAY-00
DR Genew; HGNC:6134; ITGAI.
DR MIM; 192968; .
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; NAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00357; FG-GAP 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR SMART; SM00191; Int_alpha; 5.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure.
FT DOMAIN 1 113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 114 1136 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1137 1151 FG-GAP 1.
FT REPEAT 16 75 FG-GAP 2.
FT REPEAT ? ? ?
FT DOMAIN 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GFFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 800 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
```

```
FT CARBOHYD 946 946 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 683F3C1AABF52808 CRC64;

Query Match
Best Local Similarity 18.6%; Score 1108.5; DB 1; Length 1151;
Matches 346; Conservative 210; Mismatches 486; Indels 195; Gaps 44;

QY 17 FNLDTENAMTFOENARG-FCQSVVQL-...QGRVVVVGAPQOEIVAAQNGSLVQCDYSTGS 72
DB 1 FNVVDKNSMTFGPVEDMFGYTVQOYENEEGRKWLIGSLVGPQPNRTGDVYKCPVGRGE 60
QY 73 CEP-IRLOVPVEA-...-VNMSGLSLAATSPPOLACAGPTVHQTGSENYVVKGL 122
DB 61 SLPCVKLDLPVNTSPNTEVEKNMTFGSL-VTNPNGGFLACGLPYAYRCOHLHYTGI 119
QY 123 CFLFSGNLRQOPKPEALRGCPQBDSDIAFLIDGSGIIPHDFRMKEFVSTVMEQLK- 181
DB 120 CSDVSPTEQVNSIAP-VOECSTQ-LDIVILDGNSIYPWD--SVTAFNLNLLKRWMD 174
QY 182 -KSKTLFSLMYSSEFRHFTFKFQNNPNRSLVKPITOLLGR-THATGIRKVVREL 239
DB 175 GPKOTQVGIVQGENVTEFNLYNKYSSTEEVLVAAKKIVQGRGRTALGTDTARKEAF 234
QY 240 NITNGARKNARKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFR-...SE 294
DB 235 TEARGARKGVKVMVITDGEH-DNHLKVKIQDCEDENIQRFSGIALGYSNRGNLSTE 293
QY 295 KSRQELNTIASPPRDHVPQVNFALQIOLREKIFAEGTQTGSSSEHEMSOEG 354
DB 294 KFVEIKSIASEPTEKGFNVSDALVITVITLGERIFALPATAQDAASEMMSQTG 353
QY 355 FSAATISNGPLSTVGSYDAGGVFLYTSKE-...KSTF-INMTRVDSMDNDAYLGAA 408
DB 354 FSAHYSQDWMVLGAVGDMGTGVVMQKASQIIIPRNTTFNVESTKKEPL-ASYLGYTV 412
QY 409 AILNRRVQSL-VLGAAPRQOHLGVAMPQNTGMESNANVKGTOIGAYFGASLCSVDV 467
DB 413 NSATASSGDLVLIAGQRYNHTGQVVIYRMEDGNIKILQTSGEQIGSYFSLTTTDD 472
QY 468 SNGSTDVLIGAPHY-...YQTR-GQVSVCPVPRGARAWQCDVLY 511
DB 473 KDSNTDILLVGPMPWGTGKERQGVYVYALNQTRFEYQMSLEPIKQTCSSRQNSCTT 532
QY 512 GEGQOPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSIGSPSHQ 570
DB 533 ENKNEPCGARFCTAIAAVKDLNLDGFNDIVICAPLEDHGGAVYVHG-SGKTIKRYAQ 591
QY 571 RIAGSKLSPRLQYFGQSLSGGODLTMGDLVDTVGAQHVLLRQSVLVRKAIENFNR 630
DB 592 RIPSQGDGKTLKFFQSGIHGMDLNGDGLTDTYITIGLGAALFWERDVAWVKVWTFN 651
QY 631 EVARNVFCNDQVWKGKAG--EVRVCLHVQ-KSTRDLRREGIQSVVTVYDLALDSGRPH 687
DB 652 KNIQKKNCH--MGKETVCINATVCEVKLSKEDIYIADLQ----YRVLDSLRQI 704
QY 688 SRAVFNET-...KNSTRROTQVLGLTQTCETLKLQPLNCIEDPVSPVILRNFSLVGT 741
DB 705 SRSFSGTOERKVQNRNITRKSEC-...TKHSFMYLKHDFQDSVR---ITLDENLT-D 755
QY 742 PLUSAFGNRPVLAEDAQRLFTALFPFKNCGNDNICQDLSITFTFSMLDCLVVGGRPE- 800
DB 756 PENG-...PVLDDSLPNSVHEIYIPFAKDCGNKKEKICIDLSLHVATTEKDLIVRSQNDK 810
QY 801 FNVTVTVRNDGDSYRTQVTFPFLDLSVRKVTSTLQNRQSRWSRLACESASSTEVSGAL 860
DB 811 FNVSLTVKTKDSANTRIVHYSNPLVPSGLEAIOKO-...SCESN----- 853
QY 861 KSTSCSINHPIPPENSEVTFNITFDVDSKASLGN-KLLLLKANVTSENNMTRNKTKEFQLE 919
```

```
DB 854 HNITCKGVYPLRRGEMVTFKILFQFNSTSYLMENVTIYLSATSDSEBPPETLSDNVWNIS 913
QY 920 LPVKYAVVMVTSHGVSFKYLNFTASENTSRVMQHQVSN-...LGORS-----L 966
DB 914 IPVKYEVGLQYPS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSPMPPEL 972
QY 967 PISLVF-...-LVPVRLNQTIVWRPQVTFSENLSSTCHTKE-----RLPS 1007
DB 973 KLSISFPNMTSGYVPLYPTGLSS-...-SENANCRPHIFEDPFSINSKQMTT 1021
QY 1008 HSDFLAELRKAPVNVCSIAVCORIQCDDIPFFGIQS-...-EFNATLK 1049
DB 1022 STD-...-HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLWKPTFKIKSYFSSNLNLTIR 1078
QY 1050 GNLSFDWYIKTSHNHLILIVSTAEILFNDSVFTLLPGQAFVRSOTETKVEFEPNPUL 1109
DB 1079 GEL-...-RSENASLVSSN-...-OKRELATQISKDGLPGRVPL 1114
QY 1110 --IVGSSVGGLLLLLALITAAALYKLGFRQYKDMWSE 1144
DB 1115 WVILLSAPAGLLLLLMLLILALWKIGFFKPKLKKMEK 1151

RESULT 10
ITAH HUMAN
ID ITAH HUMAN STANDARD; PRT; 1189 AA.
AC Q9UKX5; Q9UK01;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-11 precursor.
GN ITGAI1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=99417678; PubMed=1048209;
RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
RA Wang S.-X., Morris C.M., Krissansen G.W.;
RT "Cloning, sequence analysis, and chromosomal localization of the novel
RT human integrin alpha11 subunit (ITGAI1).";
RL Genomics 60:179-187(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal muscle, and Uterus;
RX MEDLINE=99395147; PubMed=10464311;
RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;
RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)
RT Integrin. A collagen-binding, i domain-containing, beta(1)-associated
RT integrin alpha-chain present in muscle tissues.";
RL J. Biol. Chem. 274:25735-25742(1999).
RN [3]
RP SEQUENCE OF 954-1188 FROM N.A.
RC TISSUE=Fibroblast;
RA Andreu N., Ebtivill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
CC HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
CC FETAL MUSCLE CELLS (IN VITRO).
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VNFA DOMAIN. INTEGRINS
```


Db 73 PVDLSTTTCEKLNQSTSMNSVTEMTKNNSLGLTLTRNVGTGGFITCGPLMAQQCSQY 132
Qy 118 YVKGCLFGLNLRQOQPOKEALRQCPQSDSDIAFLIDGSGSIIIPHDFRMEKFYSTVM 177
Db 133 YTTGVCSVSPDF-QLRTSAPAVQTCF-SFIDVVVCDSDNSIYPWD-AVKNFLEKPV 188
Qy 178 EQLK--KSKTLFSLMOYSEFRTHFTKFEQNNPNRSLVKPITOLL---GRTHATGI 231
Db 189 QGLDIGTQMGLIQYANNPRVFNLTPEKSD---EMIKATSTQFYGGDLTNTFKAI 245
Qy 232 KVVRELFTNGARKNAFLVITDGEKFGDPLGYEDVPEADREGVIRYIVG--- 287
Db 246 QYARDTAYTAAGRGPATKVMVVVTDGESH-DGSKLKAVIDQCNKDNILRFGIAGLYL 304
Qy 288 -GDAFRSEKSRQELNTTASPPRDRHVFQVNNFEALKTIONQLREKIFAIGTGTGSSSF 346
Db 305 NRNALDTNLIKELKATASPTERHPFNVSDEADLLEKAGTIGEIPFISGTVQG-GDNF 363
Qy 347 EHEMSQBGFAAIT--SNGPLLSVGSYDAGWGVLYTSKEKSTFINMT--RVDSDMN-D 401
Db 364 QMEMSQVGFSAEYSPQNNILMLGAVGAYDWSGTWQKTPHGLIFSKQAPEQILQDRNHS 423
Qy 402 AYLGAAAILLRNVQSLVIGAPRYQHIGLVAMPFRONTGHWESNAV-----KGTQIGA 455
Db 424 SYLGSVASISTGNSVHFVAGAPRANYTGQILVYSN-----ENGNVTVIQSQRGDOIGS 478
Qy 456 YFGASLCSVDVDSNGSTDVLIGAPHYVEOTR--GGQVSVCPPLRQGRARWQCDVAVLYGE 513
Db 479 YFGSVLCAVDVNDKTTIDVLLVGAPMYMNDLKKEGRVLYFTTKG-ILNWH--QFLEGP 535
Qy 514 QGQPMWFGAALTVLGVGNDKLTDAIGAPEGEDNRGAVYLFHTSGSGISPSHSQRIA 573
Db 536 NGLNARFGSAIALSDINMDGFNDVIVGSPLENQNSGAVVYNGHEGM-IRLRYSQKIL 594
Qy 574 GS--KLSPRIYFCQISGGODLTMDGLVLTCAQGHVLLRSQPVLRVKAINEFPRE 631
Db 595 GSDRAFSLHLYQFGRSLDGYDLNGDSITDVSQVAFQVQVQLWSQSADVSVDASFPPK 654
Qy 632 VARNVFCNDQVVKGEAGEVRVCLHVQKSRDLREGIQSVVYTDLALD---SGRPH 687
Db 655 I--TLNKNAEI-----KLKLCF---SAKFRPTQNNQVAIVNITDEDQFSRVI 701
Qy 688 SRVFNETKSTRQTOVLGHTQTC--TLKLQPLNCIEDPVSPIVLRNPSL--VOTPL 743
Db 702 SRGLFKENNERCLQKTMIVSQARCSEYIIHQPS---DIISPLNLCMNISLENPGT-- 756
Qy 744 SAFGNLRVLAEADQRLFTALFPPEKNCNDNIQDDLSITF---SFMSLDCVLVGGPR 799
Db 757 -----NPALRAYSETVXFSIPPHKCGDDGVCISDLVLNVQQLPATQOQPFIVSNQK 810
Qy 800 EFNVTVTRNDGEDSYRFOVTFPPFLDLSYRKVSTLQNRQSRWRLACESASST-EVSG 858
Db 811 RLTSVQLKNKEAYNTEIVDSENLFF-----ASWMPVDGTEVTCQIAS 858
Qy 859 ALKSTCSINHPFIPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTBFQL 918
Db 859 SOKSVTCNVGYPALKSKQVTFITNFDNLQ-NLQNASISFRALSSEBENMADNSVNL 917
Qy 919 ELPKYAVYVMTVSHGVSTKYLNTFASENTSRVMOHQVQSNLQOR-----SLPI 968
Db 918 KSLLYDAEIHIT-RSTNINFYVSLDGNVSSV-HSFE--DIGPKFISIKVTGSPV 973
Qy 969 SLVFLVRLNQTWIDRPQVTFSEN---LSSTCHTKE-----RLPSHS 1009
Db 974 SMA-----SVIIHPQTQKNPLMYLTGHTDQAGDISCEAEINPLIKQTSSV 1024
Qy 1010 DFLAE-LRKAPVNCIAVQRIQCDIPFGIQEEFNATLKNLSFDWIKYIKTSHNHLIV 1068
Db 1025 SFKSENPRHKLNCRTASCNIMCWLRLDQVKGDFLVNSTRVWNGTFAASTFTQVLT 1084
Qy 1069 STAEI-LFNDSVFTL-----LPGQAFVRQSTETKVEPFE-VNPLPLIVGSSVGLLL 1120
Db 1085 AAAEIDTNPQIVYIEENTVTIP-----LTIMPKHEKVEPTGTVIGSVIAGILL 1134

Qy 1121 LALITAALYKLGFFKQYKQYKDM 1141
Db 1135 LLALVAILLWKLGGFFKRYKYM 1155

RESULT 12

IT2A_MOUSE
ID IT2A_MOUSE STANDARD; PRT: 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE InterPro: alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=94355691; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding."
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development."
RL Dev. Dyn. 199:332-344(1994).
CC -1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z29987; CAA82877.1; -;
CC EMBL; X75427; CAA53178.1; -;
CC PIR; S44142; S44142.
CC HSP; P17301; IAOX.
CC MGD; MGI:96600; Itga2.
CC InterPro: IPR004113; Integrin_alpha.
CC InterPro: IPR02035; VWFA.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; P500242; INTEGRIN_ALPHA; 1.
 DR PROSITE; P50234; VWA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Repeat; Calcium; Magnesium.
 FT SIGNAL 1 26
 FT CHAIN 27 1178
 FT DOMAIN 27 1129
 FT TRANSMEM 1130 1151
 FT DOMAIN 1152 1178
 FT REPEAT 42 100
 FT REPEAT 42 100
 FT REPEAT 185 375
 FT DOMAIN 27 1129
 FT REPEAT 431 483
 FT REPEAT 485 546
 FT REPEAT 548 607
 FT REPEAT 612 664
 FT CA_BIND 496 504
 FT CA_BIND 560 568
 FT CA_BIND 624 632
 FT SITE 480 482
 FT SITE 1154 1158
 FT DISULFID 80 89
 FT DISULFID 677 734
 FT DISULFID 786 792
 FT DISULFID 862 873
 FT DISULFID 1016 1047
 FT DISULFID 1052 1057
 FT CARBOHYD 102 102
 FT CARBOHYD 109 109
 FT CARBOHYD 429 429
 FT CARBOHYD 457 457
 FT CARBOHYD 472 472
 FT CARBOHYD 696 696
 FT CARBOHYD 1054 1054
 FT CARBOHYD 1071 1071
 FT CARBOHYD 1078 1078
 SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;

Query Match 18.3%; Score 1089; DB 1; Length 1178;
 Best Local Similarity 28.1%; Pred. No. 1.3e-65;
 Matches 348; Conservative 209; Mismatches 492; Indels 188; Gaps 44;

QY 5 VLLLTALTLCGPNLOTENAMTFQ-ENARGFGOSVVLQ---QSGRVVVGAPQEIIVANQR 60
 DB 15 LMLVQQLNCLATNVLGPKAKIFSGSPSQFGVSVQVLTNPQGNWLLVSGSPGFPENRM 74
 QY 61 GSLYQC--DYSTGCEPRLQ-----VPVEANMSLGLSLAATTSPPQLLACGPTVHQ 111
 DB 75 GDVYKCPVDLPATACEKINLQNSASISNVTETKNSLGLTLTRNPGTGGFLTCGLPAH 134
 QY 112 TCGNTYVKGCLPFGSNLRQQPQ---KPEALRGCPQSDSDIAFLIDGSGSIIPHDFRR 168
 DB 135 QCGNQYVATGIC---SDVSPDFQLTSPFAVQACPSL-VDVVVCDENSIYIP--WEA 187
 QY 169 MKEFVSTWELQK--KSKTLFSLMOYSEERIHFTKEFQNNPNPSLVKPIQLLG-RT 225
 DB 188 VKNFLVKFVTGLDIPKTKQVALIOYANEPRIIFNLNDFETKEDMVQVATSETRQHGDDLT 247
 QY 226 HTATGIRKVVRELNITNGARKNAFKILVVITDGEKFGDPLGVEDVPEADREGVRYVI 285
 DB 248 NTRALEPARDAYAYSOTSGRPGATKVMVVVTGESH-DGSKLKTIVQQCNDEILRFGI 306
 QY 286 GV-----GDAFSEKSRQELNTIASPPRDHVFQVNNFEALKTIONQLREKIFAISGTQT 340
 DB 307 AVLGILNRNALDTKNLKEIKASTPTERYFPNVADAALEKAGTGEQIFSIESTVQ 366
 QY 341 GSSSSFEHMSQGFSA--AITSNGPLSTVGSYDAGVFLYTSKEKSTFFNMT--RVD 396
 DB 367 G-GDNFMQMAQVGFSGADYAPQNDILMLGAVGAFDWSGLTVQETSHKPIVFPKQAFDQVL 425

QY 397 SDMN-DAYLGAAAAIILNRVQSLVLGAPRQHQHIGLVAMFRONTQWESNANV-----KGT 451
 DB 426 QDRNHSSFLGYSVAIAISTEDGVHVFAGAPRANYTCQIVLYSNK---QGNVTVIQSHRGD 482
 QY 452 QIGAYFGASLCSVDVDSNGSDTLVLIGAPHYVEQTR--GGQVSVCPPLRGQRRQWQCDV 509
 DB 483 QIGSTFGSVLCSVDVDDKTIITDVLVLPAPTWNLDKKEEGVYLTITKGLNQHQ---F 539
 QY 510 LYGEQGPWGFAGALTVLGDVNGDKLTVAIGAPEGEDNRGAVLPHGTSGSGISPSHS 569
 DB 540 LEGPEGTNARFGSAIALSDINMDGFNDVIGSPVENENSGAVVIYNGHQT-IRTKYS 598
 QY 570 QRIAGSKLSPR--LQYFGQSLSGGQDLTMDGLVLTVAQAQGHVLLRSPVLVRVKAIMEF 627
 DB 599 QKILSGNAGFRRLHLOFFRGLDYGDLNGDSITDVSIGALGQVQLWQSQSIADVAIEALF 658
 QY 628 NPREVARNVFECDQOVKGEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDLALO---S 683
 DB 659 TP-----DKITLLNKDAKITLCLCFRAEFPAGQNNQV---ALLFNMTLDDGHS 705
 QY 684 GRPHSRVFNETKNSRRQTOVLGLTQTCET--LKLQLPNCIEDPVSPIVLRNLSLVGT 741
 DB 706 SRVTSRGVFRNSERFLQKNMVNEVQKSEHHISIQKPS---DVNPLDLRVDLSLENP 762
 QY 742 PLSAFNGNRPVLAEDAQRLFTALPFPEKNCNDNICQDDLSI-----TFSPMSLDC 792
 DB 763 GTS-----PALEYSETVKVFSIPFYKECGSDGICISDLILDVQQLPAIQTSF----- 811
 QY 793 LVVGGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWLACESAS 852
 DB 812 IVSNQNKRLTFSVLKNGESAINTVLAEFSENLF-----ASFMPVDGTE 859
 QY 853 ST-EVSGALKSTSCSINHPIFPENSEVFNTIFDVSFKASLGKLLKANVTSENMPRT 911
 DB 860 VTCEVSGSQSVTCVGVPAKSEQOVTTFINDFNLQ-NLQQAIAINFOAFSESO--ET 916
 QY 912 NKTE--FOLELPVKYAVVMTSHGVSTKYNLFTASENTRVMOHQYVSNLQOR----- 964
 DB 917 NKADNSVSLTPLYDAELHLT-RSTNINFVEISSDENAPSVIK---SVEDIGPKFIFSL 972
 QY 965 -----SLPISLVFLV-----PVLNQTVIWD 985
 DB 973 KVTAGSAPVSMALVTHIPIQYTKENPLLYLTGIDTQDQAGDICTAEINPLKLPHTA--- 1029
 QY 986 RPQVTF-SENISSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORICQDIPFFGQIEEF 1044
 DB 1030 -PSVSKFNEFR--HTKE-----LDCRTTSCSNITCWLKDLHKAEY 1068
 QY 1045 NATLKNLSFDWYIKTSHNHLIVTAEILFNDSTVFTLLPGQCAFVRSQTTETKVPPEVP 1104
 DB 1069 FINVTRVMNRFAASTFTQVLTAAAEIDTHNPQLFVIEENAVTIPLMKPTEKAEVP 1128
 QY 1105 NPLPLIVGSSVCGILLALITAAALKLGFPRQYKDM 1141
 DB 1129 T--GVIIGSIAGILLAMTAGLWKLGFPRQYKDM 1163

RESULT 13
 ITA2 HUMAN
 ID ITA2 HUMAN STANDARD; PRT; 1181 AA.
 AC P17301;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
 GN (Collagen receptor) (VLA-2 alpha chain) (CD49b).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.

FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT HELIX 292 301
FT TURN 302 303
FT STRAND 304 311
FT HELIX 313 317
FT TURN 318 319
FT HELIX 323 330
FT TURN 331 332
FT HELIX 337 340
FT STRAND 341 344
FT HELIX 347 353
FT HELIX 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match 18.3%; Score 1087.5; DB 1; Length 1181;
Best Local Similarity 27.3%; Pred. No. 1.7e-65;
Matches 339; Conservative 214; Mismatches 498; Indels 193; Gaps 44;

QY 3 LRVLTLTALT-----LCHGFNDLTENAMTFQ-ENARGFGQSIVQL- --QGSRVVVGAPQE 53
DB 11 LPLLVLALSQILNCCLAYNGLPEAKIFGSPSEQFGYAVQQPINPKGNWLLVGSWS 70
QY 54 IVAANQSGSLYQC--DYSTGSCPEIRLQ-----VPVEAVNMSLGLSAAATTSPQLLA 104
DB 71 GPPENRMGDVYKCPVDLSTACEKLNLOTSTISIPNVTETMKTNMSLGLILTRNMGTFGLT 130
QY 105 CGPTVHQTCSENYVYKICFLFGSNLRQPOKFPALRGCPQEDSDIAFLIDGSGSIIPH 164
DB 131 CGPLWAQCGNQYTTGVCSDISPDF-QLSASFSPATQPCPSL-IDVVVVCDESNSIYPW 188
QY 165 DFRMKERFVSTVMEOLK--KSKTLFSLMOYSEEFRIHETFEFQNNPNRSLVRKPIITOLL 222
DB 189 D--AVKNLEKFGVGLDITGPTKTQVGLIQYANNPRVFNLTNTYKKEEMIVATSOYSG 246
QY 223 G-RTHATGIRKVRRELFINTGARKNAFKILVITDGEKFGDPLGYEDVTPPEADREGVI 281
DB 247 GDLTTFGAIQYARKVAYSAASGRRSATKVMVVYTDESH-DGSMKAVIDOCNHDIL 305
QY 282 RYVIGV-----GDAPRSEKSRQELNTIASKPRDHVQVNNFEALKTIONLREKIPAI 336
DB 306 REGIAGVGLNRLNADLTNLIKAIKIASIPTEYFFNVSDAALLERKAGTLGEQIFSI 365
QY 337 GTQTCSSSFHEMSQEGSAITSNGP--LLSTVGSVDWAGGVFLYTSKEKSTFINMT- 393
DB 366 GTVQG-GDNFQEMSQVGSADYSSQNDILMLGAVGAFGWSGTIVQKTSHGHLIIPKQAF 424
QY 394 -RVSDDMN-DAYLGYAAAIILNRVQSLVGLAPRYCHIGLVAMFRONTGMWESNANV-- 448
DB 425 DQILQDRHSSYLGYSAVAISTGESTHPVAGAPRANYTGOIVLYSVN-----ENGNIYI 479
QY 449 ---KGTQIGAYFGASLCSYVDVDSNGSTDVLILGAPHYVEQTR--GGQVSVCPPLRGQ 503
DB 480 QAHRGDQIGSYFGSVLCSYVDVKDITDVLVVGAPWMSDLKKEGRVYVFTIKKGLGQ 539
QY 504 WQCDVILGEOQPNRGFRGAALTVLGDVNGDKLTDAVAGPCEENRGAIVLPHGTSSG 563
DB 540 HQ---FLEGPEGIENTRFGSAIALSNDIMQDFNDVIIVGSPLENQNSGAVIYHGOQT- 595
QY 564 ISPSHSQRTAGS--KLSPLQYFGOSLSGGDLTDWGLVDLTVGAQGHVLLRSOPVLVR 621
DB 596 IRTKYSQKILGSDGAFRSHLQYFGRSLOGYDGLNGDSITDVSIGAFQGVQVQWLSQSIAD 655
QY 622 KAIMEFNPREVARNYFECNDQVVGKEAGEVCLHVQKSTRDLREGIQISVVYTDIAL 681
DB 656 ALEASFTEPKI--TLVNKNAQII-----LKLCP-----SAKFRPTKQNNQVAIVYNTL 702
QY 682 D-----SGRPHSAVNETKSTRQTVLGLTQTC--ETLKLQLPNCIEDPVPVILRLN 735
DB 703 DADGFSRVTSGLFKENNERCLQKNMVVNOAQSCPEHIYIQEPS--DVVNSLDLRVD 759

QY 736 FSLVGTPLSAGNLRPVLAEDAQRLLFTALPFPFKNCGNDNICQDDLSITF-----SPMSLD 791
DB 760 ISLENPGTS-----PALEAYSETAKVFSIPFPHKDCGEDGLCISDLVLRQIPAAQEQP 813
QY 792 CLVVGGRPRENVTVTVRNDGEDSVRTQVTFPFFPLDLSYRKVSTLQNRQSRQSWRLACESA 851
DB 814 FIVSNQKRLTFTSVTLKKNRESAYNTGIVVDFSENLF-----ASFSLPVDGT 861
QY 852 SST-EVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMMPR 910
DB 862 EVTCQVAASQKSVACDVGYPAKREQQVTFTFNFDNLQ-NLQNAQSLSQALSESQEN 920
QY 911 TNKTEFOLELPVKYAVYVMTVSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQOR----- 964
DB 921 KADNLVNLKIPLLYDAEI-----HLTRSTNINFEISSDGNVPSTVSHFEDVGPKFIFSLK 976
QY 965 ---SLPISLVFLV-----PVRLNOTVIWDR 986
DB 977 VTTGSPVSPVATVIIHIPOYTKENKPNMPLMYLTGVQTDKAGDISCNADINPLKIGOT-----S 1032
QY 987 PQVTF-SENLSSTCHTKERLPSPSHSDFLAELRKAPVNVNCISIAVCQRIQCDDIPFFGIQEEFN 1045
DB 1033 SSVSFKSENER---HTKE-----LNCRTASCNSVTCMLKDVHMKGEYF 1072
QY 1046 ATLKGNLSFDWYIKTSHNHLIVSTABI-LFNDVSFTLLPQGGAFVRSQTTKVEPEPVP 1104
DB 1073 VNVTTTRWNGTFASSTFTQVLTAAAEINTYNPETIYI-----EDNTVTIPLMIM 1122
QY 1105 NP-----LP--LIVGSSVGGLLILLALITAAALYKLGFFKRYKDM 1141
DB 1123 KPDEKAEVPTGVIIIGSIAGILLALLAILKLGFFKRYKYM 1166

RESULT 14
ITAG_HUMAN STANDARD; PRT; 1167 AA.
AC O75578; Q9UHZ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-10 precursor.
GN ITGA10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular chondrocytes;
RX MEDLINE=98352078; PubMed=9685391;
RA Camper L., Hellman U., Lundgren-Aakerlund E.;
RT "Isolation, cloning, and sequence analysis of the integrin subunit
alpha10, a beta1-associated collagen binding integrin expressed on
chondrocytes.";
RL J. Biol. Chem. 273:20383-20389(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells, and Heart;
RX MEDLINE=20169197; PubMed=10702680;
RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
RA Wang S.-X., Langley R., Krissansen G.W.;
RT "The integrin alpha10 subunit: expression pattern, partial gene
structure, and chromosomal localization.";
RL Cytogenet. Cell Genet. 87:238-244(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSIONS IN
CC MUSCLE AND HEART. FOUND IN ARTICULAR CARTILAGE.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.


```

QY 92 SENLSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIPFFGIQBEFNATLKN 1051
DB 1009 TNNASCIQVNLTEPPCPVPHPELQHTNRLNGSNQCQVVRCHLQGLAKGTSEVSGLLRL 1069
QY 1052 LSFWDVWIKTSHNHLIVSTAEILFVDSVFTLLPGQAFVRSQTEKVPPEFVNPPLIV 1111
DB 1069 VNEEFRAKPKSLIVSTFELGTSEGVQLQTEASRWSLESLEV-VOTRPLILSLWILI 1127
QY 1112 GSSVGLLLALITAAVYKLGFF-----KQYK 1139
DB 1128 GSVLGGLLALLVCLWKLGFPAHKKIPEBEKREK 1164

RESULT 15
ITAL RAT STANDARD; PRT; 1180 AA.
AC P18614;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
DE (CD49a).
GN ITGA1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1].
RP SEQUENCE FROM N.A.
RX MEDLINE=90338125; PubMed=2380249;
RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
RA Esch F., Carbonetto S., Reichardt L.F.;
RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
RT for laminin and collagen."
RL J. Cell Biol. 111:703-720(1990).
RN [2].
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
RX MEDLINE=99313197; PubMed=10386626;
RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliarsky V.,
RA Gotwals P.J., Karpusas M.;
RT "Crystal structure of the alphabeta1 integrin I-domain: insights into
RT integrin I-domain function";
RL FEBS Lett. 452:379-385(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
CC E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL X52140; CAA36384.1;
DR PIR A35854; A35854.
DR PDB 1CK4; 03-MAY-00.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF A.
DR Pfam PF01839; FG-GAP; 3.
DR Pfam PF00357; integrin_A; 1.
DR Pfam PF00092; vwa; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR PRINTS: PRO0453; VWFADOMAIN.

```

```

DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium; 3D-structure.
FT SIGNAL 1 28
FT CHAIN 29 1180 INTEGRIN ALPHA-1.
FT DOMAIN 29 1142 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1143 1165 POTENTIAL.
FT DOMAIN 1166 1180 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 103 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 175 388 VWFA.
FT REPEAT 377 432 FG-GAP 3.
FT REPEAT 433 484 FG-GAP 4.
FT REPEAT 485 565 FG-GAP 5.
FT REPEAT 567 626 FG-GAP 6.
FT REPEAT 629 681 FG-GAP 7.
FT REPEAT 682 755 POTENTIAL.
FT CA_BIND 579 587 POTENTIAL.
FT CA_BIND 641 649 POTENTIAL.
FT SITE 1168 1172 GFGR MOTIF.
FT DISULFID 82 92 BY SIMILARITY.
FT DISULFID 687 696 BY SIMILARITY.
FT DISULFID 702 755 BY SIMILARITY.
FT DISULFID 807 813 BY SIMILARITY.
FT DISULFID 877 885 BY SIMILARITY.
FT DISULFID 1029 1062 BY SIMILARITY.
FT DISULFID 1066 1073 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 779 779 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 820 820 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 907 907 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 973 973 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1084 1084 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1103 1103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1180 AA; 130808 MW; 8E5DA2BE02362EE4 CRC64;

```

Query Match 18.2%; Score 1082; DB 1; Length 1180;

Best Local Similarity 27.7%; Pred No. 4e-65;

Matches 351; Conservative 199; Mismatches 484; Indels 234; Gaps 47;

QY 6 LLLTALTLCGHNLDENAMTFQENARG-FQGSVVQL---QGSRYVVGAPQEIIVANQRG 61

DB 18 LLTVLGFCSFNVVDVKMSFSFGPVEDMFQYTVQYENEEGKWLIGSLPLVGQPKARTG 77

QY 62 SLYQCDYSTGSCPE-IRLQVPVEA-----VMSLGLSLAATSPQLLACQPTVHQ 111

DB 78 DWYKCPVGRERAMPCKVLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNNGFLACPLVAY 136

QY 112 TCSENTYVKGICFLFGSLRQOPKQFPEALGCPQEDSDIAFLIDGSGSIIPHDFRMKE 171

DB 137 RCGHLHYTTGICSDVSPFTFQVNSFAP--VQECSTQ-LDIIVIVLDGSGSIIP--WESVIA 191

QY 172 FVSTVMEQLK--KSKTLFLSLQMSYSEEFRIHFTFEFQNNPNRSLVKIPITOLLG-RHTA 228

Db 192 FLNDLLKRMIDIGPKQTVQVIGVQGENVTHEFNLYKSSSTEVEVLVAANKIGROGLQMTA 251
Qy 229 TGRKVVRELFNITNGARONAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGV 288
Db 252 LGIDTARKEAFTARGARRGVKVMWITDGEH-DNYRLKQVIQCEDENIQFSAIAL 310
Qy 289 DAFR-----SEKSRQELNTIASPPRDHVFQVNFPAKTIQNLREKIPAEGTQTGSS 343
Db 311 GHYNRGNLSTEKFEVEIKSIASEPTEKHFNVSDALVTVKALGERIFALEATADQSA 370
Qy 344 SFEHEMSQGFSAAITSGNPLLSITVGSYDWAGVFLYTSKESKSTFINMT--RVDSDMD 401
Db 371 ASPEMENSOTGFSAHYSQDWMILGAYDWMGTVMQKANOVIPIHNTTFFQTEPAKME 430
Qy 402 ---AYLGYAAAIILNRNVQSLVLCAPRYOHIGLVAMPONTGMWESNANVKTQIGAYFG 458
Db 431 PLASYLGYTNSATIPGDVLYIAGOPRYNHTQVVIYKMEDGNINILQTLGGEQIGSYFG 490
Qy 459 ASLCSVDVDSNGSTDVLYIGAPHY-----YEQTR-GGQVSVCPPLPRQRA 502
Db 491 SVLTIDIDKDSVTDLLVAGPMYMGTEKEEQGVVYVAVNQTRFEYQMSLEPIROTCCS 550
Qy 503 RWQCDVLYGEOGQPWG-REGAALTVLGDVNGDKLTDVAIGAPGEEDNRCANVLFHCTSG 561
Db 551 SLKDNSTCKENKNEPCGAREFGTAIAAVKDLNVDFNDWIGAPLEDDHAGAVIYHG-SG 609
Qy 562 SGISPSHSORLAGSKLPRLOQFGOSLGGDLTMDGLDVLTVGAQGHVLLLRSQPVLRV 621
Db 610 KTIREAYAQRIPSGGDKTKLFFGQSIHGEMDLNGDGLTDTVIGGLGAALFWARDVAVV 669
Qy 622 KAIMFNPREVARNVPCNDQVVKGEAG--EVRVCLHVQ-KSTRDLREGQIQSVVTYD 678
Db 670 KVTWNPFPKVNIOKKNCR---VEGKETVCINATMCPHVKLKSKEDSIYEADLQ----YR 722
Qy 679 LALDSGRPHSRVFNET-----KNSTRQTOVLGLTQTCETLKLQLPNCI----- 723
Db 723 VTLSLRQISRSFFSGQERKIQRNIITRESE-----CIRHSFYMLDK 765
Qy 724 EDPVSPIVLRNFSLVGTPLSAGFNLRPLAEDAQRULTALFPFEKNCNGNDNICQDDLSI 783
Db 766 HDPQDSVRVTLDFNLT-OPENG-----PVLDDALPNSVHEHIPPAKCGKNKERCISDITL 819
Qy 784 TFSFMSLDCLVVGPRE-FNVTVTRNDGEDSYRTVTFPPFLDLSYRKVSTLQNRQSR 842
Db 820 NVSTTEKSLLVKSHDKFNVLVKNKGD SAYNTRTVQHS PNLIFSGIEEIQD--- 875
Qy 843 SWRLACRSASSTEVSGALKSTCSINHPIPPENSEVTNITFDVDSKASLGNKLL-LKAN 901
Db 876 ----SCESN-----QNTICRVGYPFLRAGETVTFKIIQFNTSHLSENAIHLSAT 922
Qy 902 VTSENMPRTNKTFFQLELPVKYAV---YMWVTSHGVS-----KYLNFASENTSR 950
Db 923 SDSEEPLESNDNEVNSIPKVEVGLQFYSSASEHHISVAANETIPEFINST--EDIGN 980
Qy 951 VMOHQYQVSNIGORSLP---ISLVF-----LVPVRLNQTVIMD-----RP--- 987
Db 981 EINUVYTIIRKRGHPMPPELQLSISFPNLTAGDYPVLYPIG-----WSSSDNVNCRPSL 1034
Qy 988 -----QVTF-----ENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAV 1027
Db 1035 EDPFGINSKGMKTKSKSVLKRGTIQDCSTC-----GVATITCSLLP 1077
Qy 1028 CORIQCDI-----PFFGQIEEF---NATLGNLSFDWYIKTSHNHLIVSTAEILFNDS 1078
Db 1078 SDSLQVNVSLLLMKRTPF-IRAHFSLNLTLRGELK-----SENSLTLSSN----- 1123
Qy 1079 VFTLLPGQCAFVRSQTEKVPFVNPPLP--IVGSSVGGLLILALITAALYKIGFFKR 1136
Db 1124 -----RKRELAIQISKDGLPGRVPLVILLSAFAGLULLMLLILALMKIGFFKR 1172
Qy 1137 QYKMMSE 1144

Db 1173 PLKKMEK 1180

Search completed: November 25, 2003, 14:17:21
Job time : 13.3033 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:11:54 ; Search time 33.5727 Seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481B-1

Perfect score: 5956
Sequence: 1 MALRVLLITALTLCGHFNLD.....FKROYKDMSEGGPPGAEPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rv.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4459	74.9	1151	11	Q9J130
2	3976.5	66.8	1036	11	Q8CA73
3	3824.5	64.2	920	6	Q28984
4	3514	59.0	1169	4	Q81VA6
5	3352.5	56.3	1169	11	Q9QXH4
6	3282.5	55.1	1161	11	Q9QYE7
7	1545	25.9	1161	11	Q9WTV4
8	1534.5	25.8	1160	11	Q9R200
9	1419	23.8	1196	13	Q98TF1
10	1368.5	23.0	1187	13	Q98TF0
11	1361	22.9	1086	4	Q96HB1
12	1277	21.4	927	6	Q8HZV0
13	1171.5	19.7	1167	11	Q88340
14	1125.5	18.9	1167	11	Q88341
15	1068.5	17.9	1171	13	O42094
16	1049	17.6	1038	11	Q8BS01

17	1029.5	17.3	1160	6	Q8MKF4
18	1020	17.1	895	11	Q9WUF8
19	865	14.5	348	4	Q8TES5
20	852	14.3	1332	5	Q9BPQ8
21	816	13.7	205	11	Q63001
22	753.5	12.7	780	13	Q06271
23	738	12.4	823	4	Q8WY18
24	686.5	11.5	823	11	Q8CE84
25	683	11.5	1032	11	Q61989
26	651.5	10.9	257	11	Q8C270
27	643	10.8	1036	11	Q91YD5
28	641.5	10.8	1033	6	Q9BGJ3
29	619.5	10.4	1041	5	Q9UB90
30	614.5	10.3	1041	5	O76378
31	589	9.9	1034	13	Q98TT7
32	579.5	9.7	1054	5	Q9U651
33	550	9.2	1033	13	O42598
34	535.5	9.0	1016	13	Q91779
35	530	8.9	974	11	Q924W2
36	529	8.9	1073	11	O8CC06
37	526	8.8	1047	6	Q9MZD6
38	525.5	8.8	1007	6	Q9CK48
39	514.5	8.6	1034	6	O9TUN4
40	509.5	8.6	1036	6	Q9TUN6
41	506.5	8.5	1012	11	O70304
42	491.5	8.3	1049	5	Q8SY51
43	482.5	8.1	604	11	Q8BQ25
44	475	8.0	833	5	Q9BP07
45	469.5	7.9	1036	6	Q9TU44

ALIGNMENTS

RESULT 1

Q9J130	PRELIMINARY;	PRT; 1151 AA.
ID	Q9J130	
AC	Q9J130;	
DT	01-OCT-2000 (Tremblrel. 15, Created)	
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)	
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)	
DE	Integrin beta 2 alpha subunit.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]_TaxID=10116;	
RP	SEQUENCE FROM N.A.	
RA	Fathallah D.M. Sr., Zeria K. Jr.;	
RT	"Cloning of the rat CD11b cDNA sequence."	
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBSJ databases.	
DR	EMBL; AF268593; AAF81280.1; ..	
DR	HSP; F11215; IBHQ.	
DR	InterPro; IPR000413; Integrin_alpha.	
DR	InterPro; IPR002035; VWF_A.	
DR	Pfam; PF01839; FG-GAP; 5.	
DR	Pfam; PF00357; Integrin_A; 1.	
DR	Pfam; PF00092; vwa; 1.	
DR	PRINTS; PRO1185; INTEGRINA.	
DR	PRINTS; PRO0453; VWFADOMAIN.	
DR	SMART; SM00191; Int_alpha; 5.	
DR	SMART; SM00327; VWA; 1.	
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.	
DR	PROSITE; PS00234; VVFA; 1.	
SQ	SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;	

Query Match 74.9%; Score 4459; DB 11; Length 1151;

Best Local Similarity 73.5%; Pred. No. 3.3e-317;

Matches 847; Conservative 144; Mismatches 160; Indels 2; Gaps 2;

QY 1 MALRVLLITALTLCGHFNLDTENAMTFQENARGFCQSVVQLQGSVVVVCAPQEIIVANQR 60

DB 1 MTLKVLATVLTLCGHFNLDTPMTFOENASGFCQSVIQLGETRVVVAAPQEVKAVNQT 60

```
QY 61 GSYLQCDYSTGSCBPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTSENTYVK 120
Db 61 GSYLQCDYSTNRCDPIPLQVPPPEAVNMSLGLSLAATTVPPOLLACGPTVHONCKENTYV 120
QY 121 GLCFPLFSGNLRLQOQKPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEFYSTVMEQL 180
Db 121 GLCYLFGSNLLRLKQKPEALRGCPQSDSDIAFLIDGSGSINTIDFQKMEFYSTVMDQF 180
QY 181 KSKTSLFSLMOYSEERIHETFEFQNNPRSLVKEITOLLGRTHATGIRKVVRELFN 240
Db 181 QSKTSLFSLMOYSEDEFTHFTFNDKPNPDKSHVRPIQLNGTKTASGIRKVVRELFQ 240
QY 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQEL 300
Db 241 KINGARDNAKILVITDGEKFGDPLGYEDVPEAEAGIIRYVIGVGNFAPKPSRREL 300
QY 301 NTIASKPRDHVFOVNNFEALKTIONQLREKIFAIEGTQTGSSSPHEHMSQEGFSAIT 360
Db 301 DTIASKPAGDHVFOVDNFEALNTIRNLOEKIFAIEGTQTGSSSPHEHMSQEGFSAIT 360
QY 361 SNGPLSTVGSYDWAGGVFLYTSKEKSTFFINMTVRVSDMNDAYLGYAAAILLRNVOSLV 420
Db 361 SNGPLSTVGSYDFWAGGVFLYTSKDKASFINMTVRVSDMNDAYLGYASAVISNRVOSLV 420
QY 421 LGAPRYOHIGLVAMFRONTGMWESNANVKGTQICAYFGASLCSVDVDSNGSDTLVLIGAP 480
Db 421 LGAPRYOHIGLVAMFRONTGMWESNANVKGTQICAYFGASLCSVDVDSNGSDTLVLIGAP 480
QY 481 HYYEQTRGGQVSVCLPRGQARWQCDVLYGEGQPGWRFGAALTVLGVDVNGDKLTDVA 540
Db 481 HYYEQTRGGQVSVCLPRG - RARWQCEAILHGDQHPWRFGAALTVLGVDVNGDKLTDVA 539
QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIASKLSPLRYOYQSGSLSGQDGLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIASKLSPLRYOYQSGSLSGQDGLTMDGLM 599
QY 601 DLTGAGQHVLLRSQVRLVKATMEFNPREVARNVFECNDVVYKGEAGEVRVCLHVQK 660
Db 601 DLAVGAQGRLLILRAQVRLLEATMEFSPKWLARSVFACQEVVKNKDAEVRVCLQVRK 659
QY 661 STRDLREGQIQSVVTVYDLALDSCRPHSRVAFNETKSTRQOTVLGTOTCTELKQLP 720
Db 660 NTKDLREGDIQSVTVYDLALDPCRSVVRFAFFDETNGILLRIRVFGLTQKCTELKILP 719
QY 721 NCIEDPVSPIVRLNFSLVGTPLSAFNGRLRVLAEADAQRLFTALFPPEKNCNGNDICODD 780
Db 720 DCVDNSVSPITLRLNLTAVLGEPLRSRDLRPLVLAEMAQRIETAMFPPEKNCNGNDI 779
QY 781 LSITFSFMSLCLVVGGRPFNVTVVRNDCEDSVRTQVTFPFFLDLSYRKVSTLQORS 840
Db 780 LSITVSSTRVNTLVGDSRDFDVSVTLRNDGEDSYGKVTCTYPSGLSYRKVSASQNF 839
QY 841 QRSWLACASSTEVSGALKSTSCSINHPIFPENSESEVTNITFDVDSKASLGNKLLKA 900
Db 840 KKPWRVIAE - PSSSEGGQVLKSTIWDINHPIFPANSEVTFNVTVDVDSASLGNKLLKV 898
QY 901 NVTSENMPRNKTEFQLELPVKYAVVWVTSHGVSSTKYLNTASENSTRVMOHQYQVSN 960
Db 899 VVTSENMYRDTKTESQLELPVKYAIYVWVTSGESSIKYLNFTASEMTSKYIHOYQYFNN 958
QY 961 LGORSPLSLVPLVPLVRLNQTVIDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Db 959 LGORSFPVSVVFWLIPVQINKVTIWDPPQVTFSONLSSVCRTEQKSPSHSKFQDELERTPV 1018
QY 1021 VNCISAVCORIQCDIPFGIQEENATLKNLSFDWYIKTSHNHLILVSTAEILFNDVSF 1080
Db 1019 LNCISAVCKRIQCDIPSPNSKEIFENVTLQGNLLFDWYIETSHDHLILVSTAEILFND 1078
QY 1081 TLLPQCGAFVRSQETKTEPPEVNPPLPIVGVSSVGGLLALLALTAALYKLGFFKRQYKD 1140
Db 1079 ALLPQGETFVKAQETKTEPVTVNPVPLPIVGVSSVGGVLVLLALITAGLYKLGFFKRQYKD 1138
```

```
QY 1141 MMSEGGPPGABEPQ 1153
Db 1139 MMNEAGGQDGPQ 1151
RESULT 2
Q8CA73 PRELIMINARY; PRT; 1036 AA.
ID AC Q8CA73;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Integrin alpha M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
DR EMBL; AK039444; BAC30350.1; --
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BFI1E188E77 CRC64;
Query Match 66.8%; Score 3976.5; DB 11; Length 1036;
Best Local Similarity 66.8%; Pred. No. 5.7e-282;
Matches 771; Conservative 120; Mismatches 144; Indels 119; Gaps 2;
QY 1 MALRLLVLTALTLCHGNLDLTENAMTFQENARGFQSVVQLOGSRVVVGAPOEIVANOR 60
Db 1 MTLKALLVLTALACHGNLDLDEHPMTFQENAKFGQSVVQLOGSTVVVAAAPQAKAVNQT 60
QY 61 GSYLQCDYSTGSCBPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTSENTYVK 120
Db 61 GSYLQCDYSTSRCHPIPLQVPPPEAVNMSLGLSLAATTVPPOLLACGPTVHONCKENTYV 120
QY 121 GLCFPLFSGNLRLQOQKPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEFYSTVMEQL 180
Db 121 GLCYLFGSNLLRLKQKPEALRGCPQSDSDIAFLIDGSGSINTIDFQKMEFYSTVMEQF 180
QY 181 KSKTSLFSLMOYSEERIHETFEFQNNPRSLVKEITOLLGRTHATGIRKVVRELFN 240
Db 181 KSKTSLFSLMOYSEDEFTHFTFNDKPNPDKSHVRPIQLNGTKTASGIRKVVRELFH 240
QY 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQEL 300
Db 241 KTNARENAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGNFAPKPSRREL 300
QY 301 NTIASKPRDHVFOVNNFEALKTIONQLREKIFAIEGTQTGSSSPHEHMSQEGFSAIT 360
Db 301 DTIASKPAGDHVFOVDNFEALNTIRNLOEKIFAIEGTQTGSSSPHEHMSQEGFSAIT 360
QY 361 SNGPLSTVGSYDWAGGVFLYTSKEKSTFFINMTVRVSDMNDAYLGYAAAILLRNVOSLV 420
Db 361 SNGPLSTVGSYDFWAGGVFLYTSKDKVFTFNTVRVSDMNDAYLGYASAVISNRVOSLV 420
QY 421 LGAPRYOHIGLVAMFRONTGMWESNANVKGTQICAYFGASLCSVDVDSNGSDTLVLIGAP 480
Db 421 LGAPRYOHIGLVAMFRONTGMWESNANVKGTQICAYFGASLCSVDVDSNGSDTLVLIGAP 480
QY 481 HYYEQTRGGQVSVCLPRGQARWQCDVLYGEGQPGWRFGAALTVLGVDVNGDKLTDVA 540
Db 451 -----
QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORIASKLSPLRYOYQSGSLSGQDGLTMDGLV 600
Db 451 -----SQRITGAHFSPGLQYFGOSLGGKGLTMDGLM 482
```

```
QY 601 DLTGAGQHVVLLRSQVLRVKAIMEFNPVARNVFECDQVYVKGAGEVRVCLHVOK 660
Db 483 DLAVGAGQHVVLLRSQVLRVKAIMEFNPVARNVFECDQVYVKGAGEVRVCLHVOK 542
QY 661 STDRRLREGIOISVTVYDIALDSGRPHSRAVFNENKSTRRTOVLGLTQTCETLKLQIP 720
Db 543 NTKDRUREGDIOSTVTVYDIALDPGRSRIIRAFDETKNTRRRTOVLGLTQTCETLKLQIP 602
QY 721 NCTEDPVSPIVLRNLSVGTPLSAFNGNLRPVLAEADQRLFTALPFFKNCNNDNICQDD 780
Db 603 DCVDDSVSPILRLNLTAGEPLRSFGNLRPVLANDAQRFFTAMFFKNCNNDNICQDD 662
QY 781 LSITFSMCLDCLVGGPREFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNRS 840
Db 663 LSITMSAMGLDCLVGGPQDFNMSVTLRNDGSDSYRTQVTVVYPSGLSYRKDSASQNPJT 722
QY 841 QRSWRL-ACESASSTEVSGALKSTSCINHPFIPENSEVTEFNITPDVDSKASIGNKLLK 899
Db 723 KKPWFVKPRAESSSSSGHGLKSTTWNINHPFIPANSEVTEFNITPDVDSHSGFNKLLK 782
QY 900 ANVTSENNPRTNKTKEFQLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMQHOYQVS 959
Db 783 AIVASENNNSRTHKTKEFQLELPVKYAYIMVTSDESSIRYLNFTASEMSTKVIHQYQFN 842
QY 960 NLGQSLPISVLPVLRNQTVIMDRPOVTFSENLSSCHTKERLPSHSDFLAELRKAP 1019
Db 843 NLGQSLPISVLPVLRNQTVIMDRPOVTFSENLSSCHTKERLPSHSDFLAELRKAP 902
QY 1020 VNCSTAVCQRIQCDIPFGIQEENFATLKGNSLPSDWYIKTSHNHLIIVSTAEILPNDVS 1079
Db 903 VLNCSVAVCKRIQCDIPFNTQIFNVTILKNSLPSDWYIKTSHNHLIIVSTAEILPNDVS 962
QY 1080 FTLLPQCGAFVRSOTETKVEPFPVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYK 1139
Db 963 FALLPQESVRSKTEKVEPVEVNPVPLIVGSSVGLLLALITAGLYKLGFFKQYK 1022
QY 1140 DMMEGGPPGAEPQ 1153
Db 1023 DMNEAAPDAPQ 1036
RESULT 3
Q28984 PRELIMINARY; PRT; 920 AA.
ID Q28984 AC Q28984;
DT 01-NOV-1996 (T-EMBLrel. 01; Created).
DT 01-FEB-1997 (T-EMBLrel. 02; Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23; Last annotation update)
DE CD11b (Fragment).
GN CD11B.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_taxid=9623;
RN [1]
SEQUENCE FROM N.A.
RP Lee J.-K., Schock L.B., Rutherford M.S.;
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RL ENBL; U40072; AAB16869.1; -.
DR HSP; P11215; IBBQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-CAP; 4-.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWF_A; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

```
Query Match 64.2%; Score 3824.5; DB 6; Length 920;
Best Local Similarity 79.8%; Pred. No. 6.3e-271;
Matches 735; Conservative 77; Mismatches 108; Indels 1; Gaps 1;
QY 134 POKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSVMBOLEKSKTLESLMOYS 193
Db 1 POKPEALRGCPQEDSDIAFLIDGSGSIINRLODFRMKEFVSVMBOLEKSKTLESLMOYS 60
QY 194 BEFRTHFTKGFQNNPNRSLVKPITQLLGRTHATGIRKVVRELNFNTNGARKNAFKIL 253
Db 61 EDFYTHFTFNDPKRNPSPKLLVRPIQLLGRTHATGIRKVVRELNFHKSAGARENALAIL 120
QY 254 VITDGEKFGDPLGHEDVPEADREGVIRYIVGDDAFRSKSKROELNNTIASKPRDHVF 313
Db 121 VITDGEKFGDPLGHEDVPEADREGVIRYIVGDDAFRSKSKROELNNTIASKPRDHVF 180
QY 314 QVNNFEALKTIONLREKIFALETGTQSSSSFEHEMSOEGFSAITNGPLLSVTGVSVD 373
Db 181 QVNNFEAVKTIQNOQEXTFALETGTQSSSFECEMSOEGFSAITNGPLLGAVGSPD 240
QY 374 WAGGVFLYTSKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVA 433
Db 241 WAGGAFLHMPKDRVIFINTTRVDSMDNDAYLGAAVEVILRQAQSLVLGAPRYQHTGLV 300
QY 434 MFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSV 493
Db 301 MFKQNSGAWEKNAIDKSGISYFGASLCSVDVNRDGSDDLVLIGAPHYYEOTRGQVSV 360
QY 494 CPLPGQARQCDAVLXGEOGPHGRGAALTVLGVNGDKLTDAVATGAGEEDNRGAV 553
Db 361 CPLPGQ-RAKQWQCVILCCEGHPHWRGAALTALGVNGDKLTDAVATGAGEEDNRGAV 419
QY 554 YLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGCAQGHVLL 613
Db 420 YLFHGTSELGSLSPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLMDLAVCAQGHVLL 479
QY 614 RSQPLRVKAIMFNPREVARNVFECDQVYVKGAGEVRVCLHVOKSTRDLREGQIQS 673
Db 480 RSQPLRVKAIMFNPREVARNVFECDQVYVKGAGEVRVCLHVOKSTRDLREGQIQS 539
QY 674 VVYDIALDSGRPHSRAVFNENKSTRRTOVLGLTQTCETLKLQIPNCIEDPVSPILVR 733
Db 540 IITYDIALDPGRPHRAVFEETKNTRRRTOVLGLTQTCETLKLQIPNCIEDPVSPILVR 599
QY 734 LNFSLVGTPLSAFGNLRPVLAEADQRLFTALPFFKNCNNDNICODDLSITFSFMSLDC 793
Db 600 LNFSLVGTPASSFGNLRPVLAEADQRLFTALPFFKNCNNDNICODDLSITFSFMSLDC 659
QY 794 VVGPRFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNRQSRWSRLACESASS 853
Db 660 VVGPRFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNRQSRWSRLACESASS 719
QY 854 TEVSGALKSTSCINHPFIPENSEVTEFNITPDVDSKASIGNKLLKANTVSENNPRTNK 913
Db 720 TEVSGALKSTSCINHPFIPENSEVTEFNITPDVDSKASIGNKLLKANTVSENNPRTNK 779
QY 914 TEFOLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLPSLVL 973
Db 780 TEFOLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLPSLVL 839
QY 974 VPVRLNQTVIMDRPOVTFSENLSSCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQ 1033
Db 840 VPVRLNQTVIMDRPOVTFSENLSSCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQ 899
QY 1034 DIPFGIQEENFATLKGNSLPS 1054
Db 900 DIPFGIQEENFATLKGNSLPS 920
RESULT 4
Q81VA6 PRELIMINARY; PRT; 1169 AA.
ID Q81VA6
AC Q81VA6;
```



```
Db 69 CGYHTGNCPEIPISLQVPEAVNISLGLSLAAATNPWSWLLACGPTVHHTCRENIYLTGLCLF 128
QY 126 FGSNLRRQKQKPPPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFEVSTVMEQLKSKT 185
Db 129 LSSSPKQS-QNPPTAQBCPKQDDIVFLIDGSGSIISSTDFEKNLOFVKAWSQJQORPT 187
QY 186 LPSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFINITNGA 245
Db 188 RFLSMQFSDYFRVHFTFNFISSPLSLGSLVRLRGVYTTASAKHIVITELFTTQSCA 247
QY 246 RKNAPKILVITDGEFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKOELNTIAS 305
Db 248 RQDATKVLIVITDGRKQGNLSYDSVIPMAEASIIRYAIGVKAFYNEHSHKQELKAIAS 307
QY 306 KPPRDHVFQVNFPEALKTIQNLREKI FAIEGTQTCSSSFEHMSQEGFSAITNSGPL 365
Db 308 MFSHEVSVFENPDAUKDLENQKEKIFAIEGTETPSSTFEHMSQEGFSAVFTDGPV 367
QY 366 LSTVSGYDWAGGVFLYTSKESKTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVILGAPR 425
Db 368 LGAVGFSFGAGFLYPSNMRPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILGAPR 427
QY 426 YQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQ 485
Db 428 HQHTGKVIVFTQESRHRPKSEVRGTQIGSYFGASLCSVDMORDGSTDVLVIGVPHYEH 487
QY 486 TRGGQVSVCLPRGQARQCDVAVLXGEOGPWREGAALTVLGDVNGDKLTDVAIGAPG 545
Db 488 TRGGQVSVCPMP-GVGRWHCGTTLHGEQHPWREGAALTVLGDVNGDSLADVAIGAPG 546
QY 546 EEDNRGAVLTFHGTSGSGISPHSQRISAGSKLSPRIQYFGQSLSGQDITMDGLVDLTVG 605
Db 547 EENRGAVIFHGASQDIAFPSQISASQIPSRIOYFGQSLSGQDITMDGLVDLAVG 606
QY 606 AQHVLILRSQVLRVKALMERNPREVARNVFECNDQVVKXGAEVRVCLHVQKSTRDR 665
Db 607 SKGRVLLLRTPILRVSPVHFTPAEIRSVFECQVAPQEQTLSDATVCLHIESPKTQ 666
QY 666 LREGQIQSVVYDLDLSDGRPHSRVFNKTSRTRQTVLGTQTCETLKLQLPNCIED 725
Db 667 L--GOURSTVFDLADHGRSLSTRAIFKTKETRALTRVKTGLNKHCSVSKULLPACVBD 724
QY 726 PVSPIVLRNLSLVGTPLSAFAGNLRPVLAEDAQRLFTALPPPEKNGCNDNI CODDLSITF 785
Db 725 SVTPTILNLSLVGVPISLQNLQPLAVDDQTYFTASLPPEKNGADHI CODDLSVVF 784
QY 786 SFMSLDCLVVGPREPNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQ----- 836
Db 785 GFPDLKTLVVGSDLELNVDTVSNDEGDSYGTVTTLFYPVGLSFRVABEQVFLRKKEDQ 844
QY 837 --NORSQSWRLACESASSTVSGALKSTSCSINHPIFPENSEVTPNITFDVDSKASLGN 894
Db 845 QWRRQGHSLHLMCD--STPDRSQGLWSTSCSRHVIIFRGSGQMTFLVTFDVSYPKAEALGD 902
QY 895 KLILKANVTSENMPRTNKTFOLELPVKVAVVMTVSHGVSTKYLNFTASE-NTSRVMO 953
Db 903 RULLLRARVSENVPPTKTTFOLELPVKVAVYTMISSHDQTKYLNFTSEKETSVE 962
QY 954 HQYQVSNLQORSPLISLVLPVRLNQTIVMDRPQVTFPSENLSSTCHTKERLPSHSDFLA 1013
Db 963 HRFQVNNLQORDVPVSVINFWPVIELKGEAVM-TVMVSHPNQNPILTQCYRNRKAPTQFDLLT 1021
QY 1014 ELRKAPVNCSTAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYKTSNHLILVSTABI 1073
Db 1022 HMQKSVLDCSTADCLHLRCDIPSLGILDDELFLKGNLSFGWISQTLQKVKLLSEAEI 1081
QY 1074 LFNDSVFTLLPGGAFVRSOTETKVEPEVPNPLPLIVGSSVGGLLLLALITAAALYKLG 1133
Db 1082 TENTSYSQLPQGEAFPLRAQTKVLEMYKVHPVPLIVGSSVGGLLLLAITAILYKAGF 1141
QY 1134 FKROYKMMSE 1144
|||||:|
```

```
Db 1142 FKROYKEMLEE 1152
RESULT 6
QYQYE7 PRELIMINARY; PRT; 1161 AA.
AC Q9QYE7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G..
RT Gallatin W.M.;
RL "Cloning of rat alpha D. a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF013334; AAF1241.1; -.
DR HSBP; F11215; 1BHQ.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int. alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;
```

```
Query Match 55.1%; Score 3282.5; DB 11; Length 1161;
Best Local Similarity 58.1%; Pred. No. 4.7e-231;
Matches 663; Conservative 158; Mismatches 306; Indels 15; Gaps 10;

QY 5 VLLLT--ALTLCHEFNLDTENAMTFQENARGFGQSVQVQSGVAVGAPQEIIVAAANRGS 62
Db 6 VILCGWLASCHGSLNLDVEEPIVREDAASEGGTVVQFGSRLVVGAPLEAVAVNQTR 65
QY 63 LYCCDYSGSCEPIRLQVPVAVNMSLGLSLAATSPQLLACGTVHQTCSYTVVKGL 122
Db 66 LYDCAPATCMQPIVLRSPLEAVNMSLGLSLVTATNNAQLLACGTAQACVKNMYKGS 125
QY 123 CFLFGSNLRQQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFEVSTVMEQLKK 182
Db 126 CLLLGSLL-QFTQAVPASMPECPEOMDIAFLIDGSGSINORDFAQMKDFVAKLMGEFAS 184
QY 183 SKTLFSLMOYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFINIT 242
Db 185 TSTLPSLMQYSNILKTHFTTFEKNILDPQSLVDPIVQLOGLTYTATGIRTVMEELPHSK 244
QY 243 NGARKNAPKILVITDGEFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKOELNT 302
Db 245 NGRSKAKILLVITDQKYRDPLEYSVDVIPAADKAGIIRYAIGVGDAPFOEPTALKELNT 304
QY 303 IASKPPRDHVFQVNFPEALKTIQNLREKI FAIEGTQTCSSSFEHMSQEGFSAITNS 362
Db 305 IGSAPPQDHVPKVGNFALRSIQRLQKEKIFAIEGTQSRSSSFQHEMSQEGFSAITNS 364
QY 363 GPLLSTVGSYDWAGGVFLYTSKESKTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVILG 422
Db 365 GPVLGAVGFSFGAGFLYPSNMRPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILG 424
QY 423 APRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHY 482
Db 425 APRHQHTGKVIVFTQESRHRPKSEVRGTQIGSYFGASLCSVDVDRDGSITDLVIGAPHY 484
```


Db *805 SGEDAYWVRDLDFPRGLSFRKVMELQ---PHSRMPVSCCEL---TEGSSLLTKLKCVS 859
QY 869 HPIFFENSEVTNITFDVDSKASLGKLLKANVTSEN-NMPTNKTEFQELSLPVKYAVY 927
Db 860 SPIFKAGQVSVLQVNFNTLLNSWEDFVELNGTVHCENENSSLOEDNSAATHIPVLYPVN 919
QY 928 MVTSHGVSTKYLNFTASENTRVMQHOYQV---SNLQORSPLSLVFLVPVRLNQTVIM 984
Db 920 ILTKEQENSTLYISPTPKGPKTQVQHVQVRIQPSAYDHNMP-TLEALVGPV-----W 972
QY 985 DRPQ---VTSENLSS---TCHTKE-RLPSSHDFLAELRKAPVNCSTIACVQRIQCDIPF 1037
Db 973 PHSEDITVWSVQTDPLVCHSEDLKRPSE---AEQPCLEGV-----QFRCPVIF 1021
QY 1038 FGQEBFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGFVRQOTEPK 1097
Db 1022 ---RREILLQVGTVELSKEIKAS-STLSCLSSLSVSFNSKHFLHYGSKA-SEAQVLVK 1076
QY 1098 VEPFVNPPLPIVGVSSVGLLLALITAAALYKLGFFKQYKQKMM-SEGPPGASP 1152
Db 1077 VDLIHEKEMLVVVLGGVLLFLFLALYKVGFFKRNLEKMEADGGVPNGSP 1132

RESULT 8

Q9R200 PRELIMINARY; PRT; 1160 AA.
AC Q9R200;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065901; AAD25884.1; -
DR HSSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1160 AA; 128127 MW; A33C531B139F1PAD CRC64;

Query Match 25.8%; Score 1534.5; DB 11; Length 1160;
Best Local Similarity 33.9%; Pred. No. 5e-103;
Matches 405; Conservative 218; Mismatches 465; Indels 107; Gaps 38;

QY 1 MALRV---LLTALTLL---CHGFNLDTENAMTFQENA-RGFQSVVVLQGSRRVVVGAQ 52
Db 1 MSFRIAGPRLLLGLGULFAKANSYNLDTRPTQSFLAQRHFGYVLQIEDG-VVVGAP 59
QY 53 EIVAAHQRLSYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHQT 112
Db 60 E---GDNTGGLYHCRSTSSFCQPVSLH-GSNHTSKYLGMTLATDAKGSLLACDPLSRT 115
QY 113 CSNTVYKGLCLFGSNLHQQKQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKBF 172
Db 116 CDQNTYLSGLCLVFLQSLGPMQNRPAYQECMKGVLDLVFLFDGSSQLDRKDFEKILEF 175

RESULT 9
Q98TF1
ID Q98TF1
AC Q98TF1;

PRELIMINARY; PRT; 1196 AA.


```
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;

Query Match 23.0%; Score 1368.5; DB 13; Length 1187;
Best Local Similarity 31.4%; Pred. No. 7.5e-91;
Matches 379; Conservative 210; Mismatches 460; Indels 157; Gaps 43;

QY 3 LRVLTLTALTLCGHFNLDTENAMTQENARG-FGQSVVQLQ-GSR--VVGAPQEIIVAA 58
DB 18 LGLLLPWASLSEAFNIDTEHPLRNGTPEDFGVSYYQTEFGNRKQIIIVGAPLE--GN 74
QY 59 QRGSLYQCDYSTGSCSEPIRLQVP----VEAVNMSLGLSLAATTPPQLLACGPTVHQTCS 114
DB 75 SAGEMYSCTADLQSCK--RLQRPGESVRFVFGMSAAVSSAALTS-----CSPYFAHECD 126
QY 115 ENTYYVKGCLFCLFGSNLRQOPKPPALRCGCPQEDSDIAFLIDGSGSIIPHPDPRMKFVFS 174
DB 127 GNSYLVGVYQFNSSL-QAVSNFTAAAYQSCSKREVNLFVDFDGSMSKAVEFDMNKFTK 185
QY 175 TYMEQLKSKTFLSLMOYSEERIHPTRKFEQNNPNRSLVKPITQLLGRTHATGIRKV 234
DB 186 DVKKLNSIIRFAAVQFTEIRTVDFNDYQNGSABEKLME-RHMKSLTNYKAINV 244
QY 235 VRELFN-ITNGARKNAFKLVITDGEKFGDPLGYED--VIPEADREGVIRVIGVGDAF 291
DB 245 LKNLVNSVSGADPNQAKALVIID---GDFSDNDYINILNCDEQNILRIIIVGV--- 297
QY 292 REKSRQELNTASKPRPHVQVNNFEALKTIQNLREKIFAIEGTQTGSSSSSEHEMS 351
DB 298 --KVDLTTLTALAAEPKLNNTFVIOEYGLKGLLONLQKKIYNIEGSKAHRDQKELS 355
QY 352 QEGFSAATNSGPLLSTGVSYDWAGGVFLYT---SKEKSTFNMTRVDSMDMDVILGYAA 408
DB 356 QGFSVYVQEEVSGVSGNDRGALYEVMSGSKFRQTEITDPAVN---KDSYMGYST 412
QY 409 AILNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQICGAYEGASLCSVDVDS 468
DB 413 VLGMHRGVSLLSGAPRAEHTGLVLTFTKNESTWTVMRNGEQIGSYFGASLSLLDVDS 472
QY 469 NGSTDVLVLGAPHYBE-QTRG-QGVSVCPPLRGQARWQCDVAVLGEQ--GQPMWRFGAA 524
DB 473 DGDSDPFLVGAFLFYQSPRAEGRLYVYTL-----SEQYSQKTLQSTTGRFATS 521
QY 525 LTVLGDVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSISPSHS-ORTAGSKLSPLQY 583
DB 522 LASLKDNLGDLSDVAVGAPLE--NEGVMYIYLDGTHGINPEHAPQIRPARSVLPGLQ 579
QY 584 FGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSOPLVKVAIMEFNPREVARNVFECDQV 643
DB 580 FGVSLSGGQMDNDNLDPDIVTQGGIVLLNRPVMSVAQLSFNPMELISLVYFECPS- 638
QY 644 VKGKEAGEVRVCLVHVKQSRDLRREGQIQSV--VTYDLALDSGRPHSRVAFNETKNSTR- 700
DB 639 -NAFNAFLNLTSCFTVTERTS---TGSLEKLNVLNLNLDVVRGMSRGFFDPMDSST 694
QY 701 -RQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNFS-----LVGTPLSAFGLNRLVLA 755
DB 695 LQOSVLLDGSGLSCSNFSIFMLRCVADTVSPFKIRMFNSQTQMLSGNSLAVL-DIOSRTBE 753
QY 756 DAQRLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVGPGPREPNVTVVRNDCEDSY 815
DB 754 YVEVL-----FQNC-NNSNCVADKLNFSTN-DTLVENQAHPFTVLVLSIANPGDDSY 805
QY 816 RTQVTFFFLDLSYRKVSTLQNRQSRQSWRLACESASSTEVSGALKSTSCSINHPIFFEN 875
DB 816 RTQVTFFFLDLSYRKVSTLQNRQSRQSWRLACESASSTEVSGALKSTSCSINHPIFFEN 875

806 NTSIVLVHYPEGLSLSKFDIAKPSRTR-----SSCGDRDSGATNRTTCSINLPVYRSG 857
876 SEVTFNITFDV---DSKASLGNKLLKANVTSENNMPTNKTEFQLELPVKYAVVMVYTS 932
858 TTTQFLGTFRVTWKWDYDWSDRMENTITANSNNGNM---SDMSVRRSIPVQFAVELAISL 914
933 HGV-STKYNFTASENTSRRVMOHQYQVSNLQORSPLISLVFLVPVRLNQTWIMDRPQVTF 991
915 VAEDSVTYLNFSLSDRGPKPLNIITYKVNLGLKGLPVSVTLSPCQ----- 960
992 SENLSSSTCHTKERLPSHSDFLAELRKAPVNC-----SIAVCO 1029
961 TTHVTLTTPHNFSMOEVHHSFISSYHQ--IIMCLLNKHLFFSPPELSAVQTRTTGRSLWVC- 1017
1030 RIQCDIPFPGIQEEFNATLKGNL-----SFDWYIKTSHNHLILVSTABEILF 1075
1018 -----VSSISIGELFRSSV--NLMAEVLQNVKEYESKYSY-BERRDHVFNIS-AELNF 1068
1076 NDSVFTLLPG-QGAFVRSQTEKVEPFPVNPPLPIVGVSSVGGLLLLALITAAALKGLFF 1134
1069 NTSRYNQSTGLKYNPHRSOTEKVEFVPPSRMLIVCTGAVGGFPFLIILILLKCGFF 1128
1135 KRQYKD 1140
1129 KNRPD 1134

RESULT 11
Q96HB1 PRELIMINARY; PRT; 1086 AA.
ID Q96HB1;
AC Q96HB1;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008777; AA08777.1;
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0234; VWF_A; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; F6FF2546E8C632F9 CRC64;

Query Match 22.9%; Score 1361; DB 4; Length 1086;
Best Local Similarity 32.0%; Pred. No. 2.3e-90;
Matches 381; Conservative 188; Mismatches 436; Indels 184; Gaps 38;

QY 6 LLLTALTLCGHF-----NLDTENAMTQ--ENARGFGQSVVQLQGSVVVGAPQEI 54
DB 6 ITVMAMALLSGFFFPAPASSYNLVDVGARSFPPRAGRHFGRVRLQV-GNGVIVGAPGE- 63
QY 55 VAANQRGSLYQCDYSTGSCSEPIRLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCS 114
DB 64 --GNSTGSLYQCSGSGTGHCLPVTLR-GSNYTSKYLGMTLA--TDP----- 103
QY 115 ENTYYVKGCLFCLFGSNLRQOPKPPALRCGCPQEDSDIAFLIDGSGSIIPHPDPRMKFVFS 174
DB 104 -----TDGS----- 107
```

QY	175	TVMELKSKTFLSMOYSEBEFRITHFTKEFONNPNRSLVKPITQLLGRTHRTATGIRKV	234
Db	108	-----ILPAVQFSYKTEPFDSDYVKKRDPDALLXVHKMLLTNTFGALNYV	157
QY	235	VRELNTNGARKNAFKILVVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSE	294
Db	158	ATEVFEELGAPDPAKVLIIITDGE--ATDSGNIDAADK-----IIRYIIGIKGFQTK	210
QY	295	KSROELNTIASKPRDHVQVNNFEALKTIONQLREKIFALEGQTQSGSSSEFEHMSQEG	354
Db	211	ESQETLHKFAKPASEFVKILDTFEKLKDLTFELQKKIYVIEGTSKQDLTSFNNELSSSG	270
QY	355	PSAATISNGPLSTVSGVWAGGVF-LVTSKEKSTFINMTRVDSMDMDAYLGAAA-IIL	412
Db	271	ISADLSRGHVVAGAKDWAGGFLDKADLDDTFIGNEPLEVPVAGVLYGVTVWLPS	330
QY	413	RNRVQSLVGLAPRYOHLGLVAMFR--QNTGMESNANVKGTOIGAYFGASLCSVDVDSNG	470
Db	331	ROKTSLLASGAPRYOHRGVRLLFQBPQGGHWSQVTHGTQIGSYFGGELCGVDVDQDG	390
QY	471	STDLVLGAPHYEQTGGQVSVCLPRGQARWQCDV--LYGEOGQPHREGAALTVL	528
Db	391	ETELLIGAPLYFGEOGRGVFIY-----QRRQLGFEVSELOQDGPYPLGRFGEAITAL	445
QY	529	GDVNGDKLTDVAIGAPDEEDNRGAVYLFHGTSGSGISPSHSORTAGSKLSPRLQYQOSL	588
Db	446	TDINGDGLVDVAVGAPLEE--QGAVYIFNGRHG--GLSPQPSQRLEGTVLSGIGWFGRSI	502
QY	589	SGGQDLTMDGLVDLTGVAQGHVLLRSQPLVRKAIMEFNPREVARNVFECDQV--KGK	647
Db	503	HGVKLEGDGLADVAVGAEQSMIVLSRPVDMVTLMFSFPAEIPVHEVECSYSTSNKMK	562
QY	648	EAGEVRVCLHVQKSTRDLREGQIOSVVTYDLALDSRPHSRVAVNETKSTRQTOVLG	707
Db	563	EGVNIITCFQI--KSLIPQF--QGLRVANITYTLQDGHRTRRGLFPGGRHELNRNIATV-	619
QY	708	LTQTCETLKLQPLNCIEBPVPIVLRNFSL--VGTPLSAPGN-----LRPVLAEDAQ	758
Db	620	TSMCTDPSFHPVCVQDLISPINVSLNFSLWEEGTPDRQAGKDPPILRPSLHSETW	679
QY	759	RLFTALPFFKXKNCNDNI CODDLSITFSMSLDCLVWGGPREFNVTVVRNDESDSYRTQ	818
Db	680	EI-----PFKXCGEDKCEANLRVSPSPARSRALRLTAFASLSVELSLSNLEBDAYVWQ	734
QY	819	VTFPPDLSTRKUSTLQNRQSRWRLACES--ASSTEVSALKSTSCSINHIFPENS	876
Db	735	LDLHFPPLGLSPKVEML--KPHSQIPVSCCELPPEESRLLSRAL---SCNVSSPIFKAGH	788
QY	877	EVTFNITFDVDSKASLGKLLKANVTSENN---MPRTNKTFOLELPVVKYAVVMVTS	932
Db	789	SVALQMMNTLVNSWGDSVSLHANVTWNEDSOLLEDNSATTI---IPILYPINILIQD	845
QY	933	HGVSTKYLNFTASENTSRVMQHOYQV---SNLQORSLP--ISLVFLVPVRLNQTWIDRPQ	988
Db	846	QEDSTLYVSFTPKGPKIHQVHKMYQVRIQPSIHDHNIPTLEAVGVQPPSEGPITHQWS	905
QY	989	VTFSENLSSTCHTK--ERLPHSD--FLAELRKAPVNCSTAVCORIQDIPFFGIQEEF	1044
Db	906	QOMEPPV--PCHYEDLERLPDAAEPCLPALFCFV-----FROEI	945
QY	1045	NATLKGNSFDWYIKTSNHHLLIVSTAEILFNDSVFTLLPGOGAFVRVSRQETKVEPPEVP	1104
Db	946	LVQVIGTVELVGEIEAS--SMFSLCSLSISFNSSKHFLYGSNASL--AQVVMKVDVVEK	1003
QY	1105	NPLPLIVSSVGGLLLLALITAAALYKLGFFKRYQKXMMSEG--GPPGAEP	1152
Db	1004	QMLYLYVLSGIGELLLLLIFILVLYKGVFFKENLKEKMEAGRGVNGIP	1052
QY	12	RESULT	
QY	Q8HZV0		
ID	Q8HZV0	PRELIMINARY; PRT; 927 AA.	

AC	Q8HZV0;	
DT	01-MAR-2003 (TEMBLrel. 23, Created)	
DT	01-MAR-2003 (TEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003 (TEMBLrel. 23, Last annotation update)	
DE	Lymphocyte function-associated antigen 1 (Fragment).	
OS	Bos taurus (Bovine).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
OC	Bovidae; Bovinae; Bos.	
OX	NCBI_TaxID=9913;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Thumbikat P., Kannan M.S., Maheswaran S.K.;	
RT	"Sequence of the alpha subunit of bovine lymphocyte function-	
RT	associated antigen 1."	
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF40778; AAN63636.1; -	
FT	NON_TER	1
FT	NON_TER	927
SQ	SEQUENCE	927 AA; 102523 MW; 02E2CF09917081EC CRC64;
Query Match 21.4%; Score 1277; DB 6; Length 927;		
Best Local Similarity 34.8%; Pred. No. 2.5e-84;		
Matches 340; Conservative 167; Mismatches 385; Indels 86; Gaps 28;		
QY	190	MOYSEBEFRITHFTKSFONNPNRSLVKPITQLLGRTHRTATGIRKVVRELFNITNGARKNA
Db	4	VQSFYTRTEFTFLDIYIRKQDPDALLAGVKHMLLTNTPGAINVAKVEFRPDLGARPD
QY	250	FKLVLVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTIASKPPR
Db	64	TKVLIITIDGK--PPTNTTLMRPKTS-----RSLIGIKNFKTKESQEAALHOFASKPVE
QY	310	DHVPQVNNFEALKTIONQLREKIFALEGQTQSGSSSEFEHMSQEGFSAAITSNGLPLSTV
Db	117	EFVKILDTFEKLKDLTFELQKKIYVIEGTSKQDLTSFNNELSSGISADLSEGHVGVAV
QY	370	GSYDWAGGVF-LVTSKEKSTFINMTRVDSMDMDAYLGAAA-IILNRVQSLVGLGAPRYQ
Db	177	GAKDWAGGFLDLKADLKSSTFVGNELTVESRAGLYGTVTRLPSTGRMTSLLATGAPKYQ
QY	428	HIGLVAMFR--NTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQ
Db	237	HVGRVLLFQPKRGGPMSQIEIDGIGSYFGGELCGVDVDRDGETELLIIAAPLYGE
QY	486	TRGQSVSVCLPRGQARWQCDVLYGEOGQPHREGAALTVDVNGDKLTDVAIGAPG
Db	297	QRGRGVFIY--OKIQLEFQWVSELOGETGYPLGRFGAAATAALTDINGDELTAVAGAPL
QY	546	EEDNRGAVYLFHGTSGSGISPSHSORTAGSKLSPRLQYFQCSLSGGQDLTMDGLVDLTVG
Db	354	EE--QGAVYIFNGQOG--GLSPRPSQRLEGTMFSGIQMFGRSIGHVKDGLGGDLADVAVG
QY	606	AQGHVLLRSQPLVRKAIMEFNPREVARNVFECDQVVGKKEAG-EVRVCLHVQK--ST
Db	411	AEGQVILSSRPVVDIIITSVSFPAEIPVHEVECSYSTSNQKKEGVNLTVCFOVKSLIST
QY	663	RDLRREGQIOSVVTYDLALDSRPHSRVAVNETKSTRTOVOLGLTQTCETLKLQPLNC
Db	471	-----FQGHVLANLYTLQDGHRTSRGLPFGGKHKLIGNTAVTPV--KSCFVFWFHPIC
QY	723	IEDPVSPIVLRNFSL---VGTPLS--AFGNLRPLVAEDAQRFTALPFEKXKNCNDNIC
Db	526	IQDLISPINVSLSYSLWEEEGTPDRALDRDIPPIKPSHLETKEIPFEKXKNCNDNIC
QY	778	QDDLSITFSMSLDCLVVGPPREFNVTVVRNDEGDSYRTQVTFPFDLDSYKVKSTLQ
Db	586	EADLKLAFSDMRSKILRLTPSASVRLTLRNTAEDAYVQVTLSPFQGLSFRKVEIL--
QY	838	QRSQSRWLACESASSTEVSGALKSTSCSINHIFPENSEVTFNITFDVDSKASLGKLL
Db	644	-KPHSHVPVGCCELPEEAVVHS--RALSNCVSSPIFGEDSMVDIQVMNTLQKSGWGFIE

```

QY 898 LKANVTS-----ENWPRNKTETFOLELPVKYAVVWVTVSHGVSTKYLNFTASENTS 949
DB 702 LQANVSCNNESSLEDNSATTS-----IPVMPYINVLTKDQENSTLYISFTPKSPRI 754
QY 950 RVMOHQYOV-----SNLQORSLSPLSVFLVPLVRLNQTVI---WD---RPQVTFSS-ENLSST 998
DB 755 HHVRIHYQVRQPSNYDNP-PLEALVRVPRVHSEGLTHKWSIQMPEPPVNCSPRNLESP 813
QY 999 CHTKERLPSHSDFLAELRKAPVWNCIAVCORIQCDIPFFGQBEFNATLKNLSFDWYI 1058
DB 814 SDEAE-----SCSFGT--EFCRPIDF---RQELVQVNGMVELRGTI 850
QY 1059 KTSNNHLLIVSTAELFNDSVFTLLPGGQAFVRSQETETKYBPFFVNPPLPIVSSVGG 1118
DB 851 KAS-SMLSLCSLSAISFNSSKHFLHGRNASM-AQVMKVDLVYKEMLYLYVLSSIGGL 908
QY 1119 LLLALITAAALYKLGFFKR 1136
DB 909 LLLFLIFIALYKVGFFKR 926

RESULT 13
O88340 PRELIMINARY; PRT; 1167 AA.
AC O88340;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079 (1997).
DR EMBL; AF020045; AAC23662.1; -;
DR HSP; F11215; IBOH.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR SEQUENCE 1167 AA; 128970 MW; D88A2C39ACDC2AAE CRC64;

Query Match
Best Local Similarity 29.2%; Pred. No. 2e-76;
Matches 363; Conservative 213; Mismatches 464; Indels 205; Gaps 43;

QY 5 VLLLTALTCHGFNLDTEA---MTQENARGFGSQVQLQSGRVVVGAPQEIIVAAANRGS 62
DB 8 LLMASLRKPGQAFNMVDWAWVTALQPGAPAVLSLLHLDPSN-----NQTCLLVARRSS 62
QY 63 -----LYQCDSYTGSCPIRLQVPEAVNMSLGLSLAATT---SPQLLIAC-GPTVHTQC 113
DB 63 NRTAALYRCALSI-SPEIACQ-EVEHICMPKGRYQGVTVLVGNHGVLCVQVQARKR 120
QY 114 SENTVVKGLCFGLFGLNLRQPOKPEALRG-----C----- 144
DB 121 SLNSELTGACSLTLPNLQLOAQYFSDLEGLDPAHVDSDGYCRSKGSGTGEEKKSARR 180
QY 145 -----PQDSDI AFLIDGSGSII PHDFRMKEFVSTVMEQL--KKS KTLFSLMQYS 193

```

RESULT 14

```

DB 181 RRTVEEDEDGTEIAIVLDGSGSIEPSDFQAKNFISTMMRNFFKCFECFNALVOYG 240
QY 194 EEFRIHFTFEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKVAKIL 253
DB 241 AVIOTEDLOESDRDINASLAKVQSIQVQKVTKTASAMQHVLDNIFIPSRGSRKALKVM 300
QY 254 VVIPTGKFGDPLGYEDVPEADREGVIRVYVGDAFRSEKSRQELNTIASKPRDHVF 313
DB 301 VVLTDGDFGDLPLNTTVINSKQGVVRFAGVDAFKNNNTVRELKLIASDPKEAHTF 360
QY 314 QVNFPEALKTIONOLREKIFAIECTQTGSSSPHEMSQBFSAIITSNGP-LLSTVGSY 372
DB 361 KVTNYSALDGLLSKLOORIVHMEGT---VGDALQYLAQTGFSAQIILDKGOVLGTGCAF 417
QY 373 DWAGGVFLY-TSKEKSTFINMT-RVDS-DMNDAYLVGAAAAIILNRNVQSLVLGAPRYOHI 429
DB 418 NWSGALLYSTQNGRCGFLNQTKEDSRTVOYSYLGYSVLVHLKANGISYVAGAPRHLR 477
QY 430 GLVAMFRONTOMMESNA---NVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQT 486
DB 478 GAVPELRKEDR--EEDAFVRRIEGEQMGSYFGSVLCPVDIMDGTDTDFLLVAAPFYHIG 535
QY 487 RGGVSYVCPPLPRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAP-- 544
DB 536 EGRVYVYQVPE-QDASFSLAHTLSGHPGLTNSRFGFAMAAVGDINDQDKFTDVAIGAPLE 594
QY 545 ---GEEDNRGAVYLFHGTSGSGISPSHSORIASGLSPRLQYFGQSLSGGODLTMDGLV 600
DB 595 GFGAGDASGVSYIYNGHSG-GLYDPSQOIIRASSVAGSLHYFGMSVSGGLDFNGDGLA 653
QY 601 DLTVGAQGHVLLRSQPLRVKAIIMEFPNREVARNVFECNDQVVVKGEAGEVVRVCLHVQK 660
DB 654 DITVGRSDSAVVLRSRPVVDLTVSMFTTP-----DALPMVFIGKM--DVNLCFEVD 703
QY 661 S---TRRLREGQIQSVVTVYDLALDSCRPHSRVAFVNETKSTRTQTVGLGTQTC----- 712
DB 704 SVVASEPGLRMEFLNFTVDV-----TKQRRLQCEDSSQCSCLRWK 748
QY 713 -----ETKLQLPNCIEDPVSPVILRLNFSLVGTPLSAFGNLR---PVLAE 756
DB 749 GGSFLCHFWLISTEEL---CEEDCFSNITIKTYE-----FQTSGRRRYPNPTL--D 797
QY 757 AQRILFTALF--PFEKNCNGDNICODDISITFSFMSLDCLVVGGPREFNVTVTVRNDGEDS 814
DB 798 HYKEPSAIFQLPYEKDCKNKVFCIAEIQLTN-ISQOELVVGTVKVTMNSILTSNGS 856
QY 815 YRTQVTFPFLDLSYRKVSTLQNSORSRWLACESASSTVSGALKSTSCSINHPIPE 874
DB 857 YMTNMALNYPRLNQFKKI-----QKVPSPDVQCDPKPV---ASVLVMNCKIGHPI-L-K 906
QY 875 NSEVTFNITFDVDSKASLGNKLLKANVTSNNMPRTNKTETFOLELPVKYAVVWVTVSHG 934
DB 907 RSSVNVSVTWOLEESVFPNRTADITVTISNNEKSLARETR---SLQPRHAFIAVLSR-- 961
QY 935 VSTKYLNPFTASENTRVMQHOYQVSNLQGRSLPISLVFLVPLVRLNQTIVMDRPQVTFSEN 994
DB 962 PSVMYMN--TSQSPDHKHEFFNVHGENLFGAVFQLQICVPIKLQDF-----QIVRVKN 1013
QY 995 LSST-----CHTKERLPSHSDFLAELRKAPVWNCIAVCORIQCDIPFFGQBEFNATLK 1049
DB 1014 LTKTDHTECTQSQEPACGSDPVQHVKEWHVWCAI-----TSNK 1053
QY 1050 GNLSFDWYIKTSHNHLIVSTA-----EILFNDSVFTLLPGQAFVRSQETETKYEPF-- 1101
DB 1054 ENVTVAEISVCHTKQLLRDVSSELPILGEISFNKSLYEGLNAE-----NHRKTIVFL 1107
QY 1102 --EVPNPLPIVSSVGGLLLLALITAAALYKLGFFKRYKQYKDMWSE 1144
DB 1108 KEEETRSLPLIIIGSSIGGLLVVLIALLFKCGFFKRYKQYQNLLE 1152

```


RA "Smooth muscle cell phenotype-dependent transcriptional regulation of
RT the 'alpha' integrin gene."
RL J. Biol. Chem. 272:26643-26651 (1997).
DR EMBL; AB000470; BAA23160.1
DR EMBL; AB000471; BAA23161.1
DR HSSP; P17301; IAOX.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5
DR Pfam; PF03357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;
Query Match 17.9%; Score 1068.5; DB 13; Length 1171;
Best Local Similarity 28.4%; Pred. No. 6.9e-69;
Matches 349; Conservative 215; Mismatches 521; Indels 145; Gaps 47;
QY 1 MALRVLTLAL-TLCHGFNLDTENAMTFOENARG-FGQSVVQL---QGSRVVVGAPQEI 55
Db 1 MAASVWMLTALFHLIGASVDVKNAITSGPLEDMFGVTVOQYENEEKWVLGSLVQ 60
QY 56 AANQSGSYQCDYSTGSGCEP-IRLOVP-----VEAV--NMSLGLSLAATTPPQLLAC 105
Db 61 PEKRTGDVYKCPVGRDSDSPICLNLPDATSPVNVMEVKNMTLGTTL-VTNPKGGFLAC 119
QY 106 GPTVHQTCTSENYVKGCLFELFGSNLRQOPQKPEALRCCPDSDIAFLIDGSGSIIPHD 165
Db 120 GPLYAYKCGRLHYTTGVCNSVSTFETVKAVAP-SVQEC-RTQDLDIVLDGNSIYP-- 175
QY 166 FRMKPEFVSTVMEQLK--KSKTLFSLMOYSEEFRIHFTFKGFONNPNRSLVKPITQLG 223
Db 176 WESVTAFLNSLLRNWDIGPQOTQVGIQVQGVTVVHEFVLYNTSYTTEVWDAALRIQRGG 235
QY 224 -RTHATGIRKVVRELFNITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIR 282
Db 236 TOTMTALGIDTAREAEFTAHGARGGVOKVMVIVTDGESH-DNYRLQEVIDKCEDENIQ 294
QY 283 YVIGVDAP-----RSEKSRQELNTIASKPRDRHVQVNNFEALKTIONQREKIFAIEG 337
Db 295 FAIAILGYSRGNLSTKRPVEEIKSIASKPTEKHFNFVSDLAALVTIYEALGERIFALEA 354
QY 338 TOTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAAGGVFLYTSKEKSTFINMTRVDS 397
Db 355 TTDQQAASFEMMSQAGFSAHYSQWVLMGAVGAYDMGTVMVWKDSDISIPSDNTPRDR 414
QY 398 DMND-----AYLGYAAAILNRVOSLVLGAPRYOHIGLVAMFRONTGMWESNANVKGTQ 452
Db 415 HSEKIEPLAAVIGTVNSALTPGGVLYIAGQPRYNTGQVLIYKMEGREVKVQLQKGEQ 474
QY 453 IGATFGASLCSVDVDSNGSTDLVLIGAPHYY--EOTRGQVSVCPPLRG-----ORA 502
Db 475 IGSYFGGVITIDINRDSFTDLLLVGAPTYMGTEKEEQGVVYVALNKTKEFYQMSLEPI 534
QY 503 RWQCDVLYGE-----QGPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Db 535 KOTCCSPUKHDTCKVLKNKEPCAREGTAIAAVKDLNLDGYNDIVIGSPLEDDHHRGAVIY 594
QY 557 HGTSGSGISPSHSQRIAGSKLSPLRYQFQSLSGGQDLTMDGLVDLTVGAQCHVLLLRQ 616
Db 595 HG-HGNTISKXYTORIASGSGDEKVKFFGQSVHGMENDLDDGLDVTIGLGGALFWSR 653
QY 617 PVLRVKAIAMERNPREVARNFECNDQVVKGEAGEVRVCLH--VOKSTRDLRREGQIOSV 674
Db 654 DVAEVNWSMQFMPKSI--NIQQNCQINKRK-----TICINATICFKTRLSKEDIFESS 706

QY 675 VTYDLALDSGRPHSRVNFNETKNSRRTQVGLGTQTCETLK-----LQLPNCIEDPVSP 729
Db 707 LQYWITLDAQRIQSRLSFTETH---ERKMOKNITIKGSECIKHNFYMLDKPD-FQDSVK- 761
QY 730 IVLRNFSLVGTPLSAFNGLRPVLAEDAQRLFTALFPFEKXCGNDNICQDDLSITFSF-- 787
Db 762 VLLEFNFS---DPESG-----PVLDSNLPSNISSEYIPFTKDCGAKNKCISDLALNVKASI 813
QY 788 ---MSLDCLVVGPRPREFNVTVTVNRDGEDSYRTQVTPFFPLDLSYRKVSTLQNSORSW 844
Db 814 AGDSSSPFIYKSNDRFTIQLSINKKDSAYNTRALVOYSPNIIIFAGIEDIQKD----- 867
QY 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTS 904
Db 868 --SCESN-----HNITCKVGYPLKPEEEISFKISFQFNASYLEN-ATVDVYATS 915
QY 905 ENMWP-----RTNKTPEOLELPVKYAVVMVTS-----HGVSTKYLNFTASENSTRVMOHQ 955
Db 916 DSEEPPELTRDNRG--QVTIPVKYEVGLIFVSVFKEHHVIAAANETIPTAINTTEQIGDE 973
QY 956 ---YOVSNLQORSLP-ISLVFLVP--VRLNQTIVMDRPOVTFSENLSSTCHTKE----- 1003
Db 974 VTLHYRIEKGHEHPMPNLTLOLLYPDVTAKNTLLY---ITTLHSQNAICKSSYPVDHL 1030
QY 1004 RLPSSHDFLAELRKAP---VVMCSIAVCQRIQCDIPFGIOBEFNATLKNLSFDW--- 1056
Db 1031 KIGSGKSYVLPKIKEPTKDTIMECDTFCASINCALAPSDI-SQVNVSLR-----VWKP 1084
QY 1057 YIKTSHNHLIIVSTAELFNDSVFTLLPGOGAFVRSOTETKVPFEPVNPPLPL--IVGSS 1114
Db 1085 IIKASIHSLTLVVKALLRSENSSLIL---RNDHQKLETMIKISKEPPPGSVPLWVPLSI 1141
QY 1115 VGGILLALITAAALYKLGFEKQYKDMSE 1144
Db 1142 FAGLLIALLIFALWKAGFFKRPPLKKKMEK 1171

Search completed: November 25, 2003, 14:20:22

Job time : 39.5727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.8086 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481B-1
Perfect score: 5956
Sequence: 1 MALRVLLTALTLCGHFNLD.....FKRQYKDMMSGGPPGAEPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 159726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5956	100.0	1153	AAW65090	Human Beta-integrin
2	5956	100.0	1153	AAW65090	Human CD11b protei
3	5956	100.0	1153	AAW65090	Human Beta2 integr
4	5956	100.0	1153	AAW65090	Human integrin 1 a
5	5956	100.0	1153	AAW65090	Integrin Mac-1 alp
6	5946	99.8	1153	AAW65090	Alpha subunit of M
7	3494	58.7	1153	AAW65090	p150.95 alpha subu
8	3480	58.4	1163	AAW65091	Human Beta-integrin
9	3480	58.4	1163	AAW65091	Human CD11c protei

10	3480	58.4	1163	23	ABG61470	Human Beta2 integr
11	3478	58.4	1163	24	ABU07406	Protein differenti
12	3455	58.0	1161	16	AAW78166	Human beta-2 integ
13	3455	58.0	1161	18	AAW23049	Human beta 2 integ
14	3455	58.0	1161	19	AAW72825	Human alpha-d. Ho
15	3455	58.0	1161	19	AAW65089	Human Beta-integrin
16	3455	58.0	1161	19	AAW57491	Human Beta2 integr
17	3455	58.0	1161	20	AAW73342	Human alpha d clon
18	3455	58.0	1161	21	AAW07359	Human alpha d clon
19	3455	58.0	1161	21	ABG61468	Human Beta2-integr
20	3439.5	57.7	1161	18	AAW23064	Human beta 2 integ
21	3439.5	57.7	1161	19	AAW72837	Human alpha-d deri
22	3439.5	57.7	1161	19	AAW65106	Human Beta-integrin
23	3439.5	57.7	1161	20	AAW73343	Human alpha had prote
24	3439.5	57.7	1161	21	AAW07376	Human alpha d prot
25	3439.5	57.7	1161	21	ABG61485	Human Beta2-integr
26	3272.5	54.9	1161	16	AAW78169	Rat alpha-d subuni
27	3270.5	54.9	1161	18	AAW23062	Rat beta 2 integr
28	3270.5	54.9	1161	19	AAW72824	Rat alpha-d #1. R
29	3270.5	54.9	1161	19	AAW60004	Rat alpha d polype
30	3270.5	54.9	1161	21	AAW07374	Rat alpha d protei
31	3270.5	54.9	1161	21	ABG61483	Rat Beta2-integrin
32	3264	54.8	1161	18	AAW23061	Mouse beta 2 integ
33	3264	54.8	1161	19	AAW72836	Mouse beta-integrin
34	3264	54.8	1161	19	AAW65103	Mouse alpha d poly
35	3264	54.8	1161	19	AAW60003	Mouse alpha had prote
36	3264	54.8	1161	20	AAW73347	Mouse alpha d prot
37	3264	54.8	1161	21	AAW07373	Mouse Beta2 integr
38	3264	54.8	1161	23	ABG61482	Rat beta-integrin
39	3263.5	54.8	1161	19	AAW65104	Rat alpha-d subu
40	3263.5	54.8	1161	20	AAW73345	Mouse alpha-d subu
41	3260	54.7	1161	16	AAW78168	Rat beta 2 integr
42	3251.5	54.6	1151	18	AAW23059	Rat alpha-d #2. R
43	3251.5	54.6	1151	19	AAW72834	Rat beta-integrin
44	3251.5	54.6	1151	19	AAW65101	Rat alpha-d polype
45	3251.5	54.6	1151	19	AAW60001	Rat alpha d polype

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; Protein; 1153 AA.

AC AAW65090;

DT 28-SEP-1998 (first entry)

DE Human Beta-integrin CD11b subunit protein.

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
lung inflammation; acute respiratory distress syndrome; CD11b subunit;
rheumatoid arthritis.

OS Homo sapiens.

PN US5728533-A.

PD 17-MAR-1998.

PF 07-JUN-1995; 95US-0485618.

PR 07-JUN-1995; 95US-0485618.

PR 23-DEC-1993; 93US-0173497.

PR 05-AUG-1994; 94US-0286889.

PR 21-DEC-1994; 94US-0362652.

PA (ICOS-) ICOS CORP.

PI Gallatin WM, Van DER VIEREN M;

DR WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using

PT immobilised or labelled alpha-d polypeptide, useful for, e.g.

PT treating type-1 diabetes

XX Example 5; Fig 1A-D; 106pp; English.

CC This sequence represents a human beta-integrin CD11b subunit which is

CC used to describe a method for identifying compounds that modulate the

CC interaction of the beta-integrin alpha-d subunit with a binding partner

CC of alpha-d which involves contacting an alpha-d polypeptide with an

CC alpha-d binding partner, one of which is immobilised and the other of

CC which is labelled, in the presence of a test compound, and determining if

CC the compound affects binding between the alpha-d polypeptide and alpha-d

CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment

CC comprising the cytoplasmic, transmembrane or extracellular domain of

CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,

CC asthma, psoriasis, lung inflammation, acute respiratory distress

CC syndrome and rheumatoid arthritis.

XX

SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 19; Length 1153;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTCHGPNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQR 60

DB 1 MALRVLLLTALTCHGPNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQR 60

QY 61 GSLVQCDYSTGSCPEIRLOVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVK 120

DB 61 GSLVQCDYSTGSCPEIRLOVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVK 120

QY 121 GLCFUFGNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQL 180

DB 121 GLCFUFGNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQL 180

QY 181 KSKTLPFLMSQVSEBFRTHFTKFEQNNPNRSLVKPITQLLGRTHATGIRKVVRELNF 240

DB 181 KSKTLPFLMSQVSEBFRTHFTKFEQNNPNRSLVKPITQLLGRTHATGIRKVVRELNF 240

QY 241 ITNGARKNAFVLVITGKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSKSQEL 300

DB 241 ITNGARKNAFVLVITGKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSKSQEL 300

QY 301 NTIASKPPRDHVFQNNFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAIT 360

DB 301 NTIASKPPRDHVFQNNFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAIT 360

QY 361 SNGPLLSTVGSYDAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIIILNRVQSLV 420

DB 361 SNGPLLSTVGSYDAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIIILNRVQSLV 420

QY 421 LGAPRYQHILVAMFRONTGMWESNANVKGTOIGAYFGASICSVDVDSNGSTDVLICAP 480

DB 421 LGAPRYQHILVAMFRONTGMWESNANVKGTOIGAYFGASICSVDVDSNGSTDVLICAP 480

QY 481 HVEYOTRGQGVSVCPPLRGQARWQCDVLYGECQCPWGRFGAALTVLGDVNGDKLTDVA 540

DB 481 HVEYOTRGQGVSVCPPLRGQARWQCDVLYGECQCPWGRFGAALTVLGDVNGDKLTDVA 540

QY 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600

DB 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600

QY 601 DLTVCAQGHVLLRSPQVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOK 660

DB 601 DLTVCAQGHVLLRSPQVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOK 660

QY 661 STRDLRREGIQSVVTVYDLALDSGRPHSAVFNETKNSRRQTQVLGLTQTCETLKLQLP 720

DB 661 STRDLRREGIQSVVTVYDLALDSGRPHSAVFNETKNSRRQTQVLGLTQTCETLKLQLP 720

QY 721 NCIEDPVSPVLRNLNFSLVGTPLSAFGNLRPVLAEADAQRLFTALFPFKXKNCNDNICODD 780

DB 721 NCIEDPVSPVLRNLNFSLVGTPLSAFGNLRPVLAEADAQRLFTALFPFKXKNCNDNICODD 780

QY 781 LSITFSFMSLDCLVWGGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRS 840

DB 781 LSITFSFMSLDCLVWGGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRS 840

QY 841 QRSWRLACSSASSTEVSGALKSTSCSINHPPIPPENSEVTFNITFDVDSKASLGNKLLKA 900

DB 841 QRSWRLACSSASSTEVSGALKSTSCSINHPPIPPENSEVTFNITFDVDSKASLGNKLLKA 900

QY 901 NVTSENNMPTNKTEQLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 960

DB 901 NVTSENNMPTNKTEQLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 960

QY 961 LGQRSPLISLVFLVPVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020

DB 961 LGQRSPLISLVFLVPVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020

QY 1021 VNCISIAVCQRIQCDIPFFGIOEEFNATLKGNSLDFWYIKTSHNHLIVSTAEILFNDVSF 1080

DB 1021 VNCISIAVCQRIQCDIPFFGIOEEFNATLKGNSLDFWYIKTSHNHLIVSTAEILFNDVSF 1080

QY 1081 TLLPGQGAFVRSQETKVEPPEVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKD 1140

DB 1081 TLLPGQGAFVRSQETKVEPPEVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKD 1140

QY 1141 MMSEGGPPGAEPPQ 1153

DB 1141 MMSEGGPPGAEPPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; Protein; 1153 AA.

XX AAB07360;

AC AAB07360;

XX 17-JAN-2001 (first entry)

XX Human CD11b protein sequence.

XX Human; macrophage infiltration inhibition; alpha_d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11b.

XX Homo sapiens.

OS WO200029446-A1.

PN 25-MAY-2000.

XX 16-NOV-1999; 99WO-US27139.

XX 16-NOV-1998; 98US-0193043.

PR 08-JUL-1999; 99US-0350259.

XX (ICOS-) ICOS CORP.

PA Gallatin MW, Van Der Vieren M;

XX WPI; 2000-387751/33.

XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

PT macrophage infiltration and reduce inflammation at central nervous

PT system injury sites

Example 5; Fig 1; 270pp; English.

Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha_d (AAA60014 and AAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha_d. The Alpha_d gene and protein may be useful in therapy for diseases linked to aberrant alpha_d function e.g. Type 1 diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

Query Sequence 1153 AA;

Query Match	100.0%;	Score 5956;	DB 21;	Length 1153;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 1153;	Conservative	0;	Mismatches	0;		
			Indels	0;		
			Gaps	0;		
Qy	1	MAURVLLLTALTLC	CHFNLD	TENAMTFOENARFGQSVVVOLOGSRVVVGAPQEI	VAANQR 60	
Db	1	MAURVLLLTALTLC	CHFNLD	TENAMTFOENARFGQSVVVOLOGSRVVVGAPQEI	VAANQR 60	
Qy	61	GSLYQCDYSTG	CEPIRLQVP	VEAVNMSLGLSLAAT	TPPQLLACGPTVHQT	CSNTYVK 120
Db	61	GSLYQCDYSTG	CEPIRLQVP	VEAVNMSLGLSLAAT	TPPQLLACGPTVHQT	CSNTYVK 120
Qy	121	GLCFLFGSNLR	QPKPEALRC	QPOEDSDIAFLIDGSGSI	IPHDFRMKEFVSTVMEQL 180	
Db	121	GLCFLFGSNLR	QPKPEALRC	QPOEDSDIAFLIDGSGSI	IPHDFRMKEFVSTVMEQL 180	
Qy	181	KKSKTLFSLMOY	SEERIHFT	PKFONNPNRSLVKP	ITOLLGRTHATGIRKVVRELFN 240	
Db	181	KKSKTLFSLMOY	SEERIHFT	PKFONNPNRSLVKP	ITOLLGRTHATGIRKVVRELFN 240	
Qy	241	ITNGARKNAFK	ILVITDGEK	FGDPLGYEDVIE	PEADREGVIRYVIGVD	AFRSEKSRQEL 300
Db	241	ITNGARKNAFK	ILVITDGEK	FGDPLGYEDVIE	PEADREGVIRYVIGVD	AFRSEKSRQEL 300
Qy	301	NTIASKPPRDH	VFOVNNFEAL	KTIQNLREKIP	AIETGTGTGSSSSPEHEMSQEGFSAIT 360	
Db	301	NTIASKPPRDH	VFOVNNFEAL	KTIQNLREKIP	AIETGTGTGSSSSPEHEMSQEGFSAIT 360	
Qy	361	SNGPLLSTVGSY	DWAGGVFLY	TSKSKSTFINMTRV	SDMNDAYLGYAAAILLNRVQSVL 420	
Db	361	SNGPLLSTVGSY	DWAGGVFLY	TSKSKSTFINMTRV	SDMNDAYLGYAAAILLNRVQSVL 420	
Qy	421	LGAPRYQHIGL	VAMFRQNTG	MWESNANVKGTQ	CAYFGASLCSDVDVDSNGSTDVL	IGAP 480
Db	421	LGAPRYQHIGL	VAMFRQNTG	MWESNANVKGTQ	CAYFGASLCSDVDVDSNGSTDVL	IGAP 480
Qy	481	HYEYQTRGGQ	SVCPPLPRGOR	ARWQCDVLYG	EQGPWGRFGAALTVLGDVNGDKL	TDVA 540
Db	481	HYEYQTRGGQ	SVCPPLPRGOR	ARWQCDVLYG	EQGPWGRFGAALTVLGDVNGDKL	TDVA 540
Qy	541	IGAPGEDNRGA	VLYFHGTSGS	ISPSHSOR	IAGSKLSPRLQYFGQSLGGQDL	TMDGLV 600
Db	541	IGAPGEDNRGA	VLYFHGTSGS	ISPSHSOR	IAGSKLSPRLQYFGQSLGGQDL	TMDGLV 600
Qy	601	DLTVGAQGHV	LLRSQPLRVK	AIMENPREVARN	VFECNDQVVKGEAGEVRVCL	HVQK 660
Db	601	DLTVGAQGHV	LLRSQPLRVK	AIMENPREVARN	VFECNDQVVKGEAGEVRVCL	HVQK 660
Qy	661	STRDLREGQIQ	SVVYTDALDS	GRPHSRAVFN	FTKNSTRQTQVLGLTQTCTETLK	QLP 720

Db	661	STRDLRREGQIQSVVYTDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTELKQLP	720
Qy	721	NCIEDPVPSPVILRLNFSLVGTPLSAFGNLRLPVLAEADAQRLFTALFPPEKNCGNDNICDD	780
Db	721	NCIEDPVPSPVILRLNFSLVGTPLSAFGNLRLPVLAEADAQRLFTALFPPEKNCGNDNICDD	780
Qy	781	LSITPSFMSLDCLVWGGPREFNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNORS	840
Db	781	LSITPSFMSLDCLVWGGPREFNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNORS	840
Qy	841	QRSMWLACESASSTEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKA	900
Db	841	QRSMWLACESASSTEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKA	900
Qy	901	NVTSENNMPRTNKTETQLELPVKYAVYVMVTSHGVSTKYLNFTASENSTRVMQHOYQVSN	960
Db	901	NVTSENNMPRTNKTETQLELPVKYAVYVMVTSHGVSTKYLNFTASENSTRVMQHOYQVSN	960
Qy	961	LQGRSLPISLVFLVPLVRLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV	1020
Db	961	LQGRSLPISLVFLVPLVRLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV	1020
Qy	1021	VNCSTAVCORIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVF	1080
Db	1021	VNCSTAVCORIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVF	1080
Qy	1081	TLLPGQGAFFVRSQTEKTVEPPEVPNPPLPLIVGSSVGGLLLLLALITAAALYKLGFFFKQYKD	1140
Db	1081	TLLPGQGAFFVRSQTEKTVEPPEVPNPPLPLIVGSSVGGLLLLLALITAAALYKLGFFFKQYKD	1140
Qy	1141	MMSEGGPPGAEPQ 1153	
Db	1141	MMSEGGPPGAEPQ 1153	
RESULT 3			
ABG61469			
ID	ABG61469 standard; Protein; 1153 AA.		
XX	AC	ABG61469;	
XX	AC	ABG61469;	
DT	27-AUG-2002 (first entry)		
XX	Human Beta2 integrin alphaCD11b subunit.		
XX	Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;		
KW	LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;		
KW	multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;		
KW	inflammatory bowel disease; Crohn's disease; ulcerative colitis;		
KW	immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;		
KW	intracellular cell adhesion molecule; vascular cell adhesion molecule;		
KW	locomotor recovery; locomotor damage; locomotor impairment;		
KW	autonomic dysfunction; sensory dysfunction; spinal cord injury.		
XX			
OS	Homo sapiens.		
XX	WO200230980-A2.		
PN	18-APR-2002.		
XX	15-OCT-2001; 2001WO-US32059.		
PD	13-OCT-2000; 2000US-0688307.		
PF	(ICOS-) ICOS CORP.		
XX	Gallatin WM, Van Der Vieren M;		
PI	WPI; 2002-463260/49.		
XX	Use of an anti-alpha-d monoclonal antibodies for promoting locomotor		
PT	recovery, inhibiting locomotor damage, limiting locomotor impairment,		
XX			

or limiting autonomic and sensory dysfunction following spinal cord injury -

Example 5; Page 191-194; 270pp; English.

The invention relates to promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury by administering an anti-alpha d (Beta2 integrin, alpha2 subunit) monoclonal antibody to a spinal cord injury victim. The method also involves the use of a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule, vascular cell adhesion molecule). The method is useful for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury. In particular, the spinal cord injury comprises compression of the spinal cord. The antibodies are also useful for reducing inflammation at the site of a central nervous system injury. The specification also details the identification of Beta2 integrin alpha d cDNAs and proteins, for use in raising the antibodies. Beta2 integrins are implicated in diseases such as LAD (leukocyte adhesion deficiency, inflammatory response, diabetes, multiple sclerosis, arthritis, graft atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, immune complex alveolitis and leukaemia. The present sequence is a Beta2 integrin alpha subunit sequence included for comparison with the Beta2 integrin alpha d protein sequences.

Query Match 100.0%; Score 5956; DB 23; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MALRVLLLTALTCHGFLNLTENAMTFOENARGFGQSVVQLQGRVVVGAPQEIIVANQR 60
1 MALRVLLLTALTCHGFLNLTENAMTFOENARGFGQSVVQLQGRVVVGAPQEIIVANQR 60
61 GSLYOCYDSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVK 120
61 GSLYOCYDSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVK 120
121 GLCLFGLNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQL 180
121 GLCLFGLNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQL 180
181 KSKTFLSLMQYSEEFRIHFTKPEFQNNPNRSLVKPITQLLGRTHHTATGIRKVVRELFN 240
181 KSKTFLSLMQYSEEFRIHFTKPEFQNNPNRSLVKPITQLLGRTHHTATGIRKVVRELFN 240
241 ITNGARKNAFKILVITDGEKFGDPLGYEDVTPEDREGVIRYVIGVGDAFRSEKSRQEL 300
241 ITNGARKNAFKILVITDGEKFGDPLGYEDVTPEDREGVIRYVIGVGDAFRSEKSRQEL 300
301 NTIASKPRDRHVQVNNFEALKTIONLREKIFAIEGTQTGSSSFEHEMSQEGFSAAIT 360
301 NTIASKPRDRHVQVNNFEALKTIONLREKIFAIEGTQTGSSSFEHEMSQEGFSAAIT 360
361 SNGPLLTSTGSDWAGGVPFLYTSKSKSTFINNTRVDSNDNDAYLGYAAAILRNVRQSLV 420
361 SNGPLLTSTGSDWAGGVPFLYTSKSKSTFINNTRVDSNDNDAYLGYAAAILRNVRQSLV 420
421 LGAPRYQHTGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
421 LGAPRYQHTGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
481 HYYEQTRGGQVQVCPPLPRQARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVA 540
481 HYYEQTRGGQVQVCPPLPRQARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVA 540
541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGGDLTMDGLV 600
541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGGDLTMDGLV 600

601 DLTVGAGQHVLRLRSQPVLRVKAIMEFNPREVARNVFECDNQVVKQKAGEVRVCLHVQK 660
601 DLTVGAGQHVLRLRSQPVLRVKAIMEFNPREVARNVFECDNQVVKQKAGEVRVCLHVQK 660
661 STRDLRREGQIQSVVWYDIALDSGRPHSRAVFNETKNSRRQTVGLTQTCTETKLQLP 720
661 STRDLRREGQIQSVVWYDIALDSGRPHSRAVFNETKNSRRQTVGLTQTCTETKLQLP 720
721 NCIEDPVSPIVLRLNLSVGTPLSAFNGLRPVLAEADAQRLFTALFFPEKNCNDNICQDD 780
721 NCIEDPVSPIVLRLNLSVGTPLSAFNGLRPVLAEADAQRLFTALFFPEKNCNDNICQDD 780
781 LSITFSFMSLDCLVWGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQVRS 840
781 LSITFSFMSLDCLVWGPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQVRS 840
841 QRSWRLACESASSTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA 900
841 QRSWRLACESASSTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA 900
901 NVTSNNMPRTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHQVSN 960
901 NVTSNNMPRTNKTETQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHQVSN 960
961 LGQRSIPISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020
961 LGQRSIPISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020
1021 VNCIAVCQRIQCDIPFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVF 1080
1021 VNCIAVCQRIQCDIPFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVF 1080
1081 TLLPGQAFVRSOTETKVEPFEVNPPLVIGSVGGLLLALITAAALYKLGFFKQYKD 1140
1081 TLLPGQAFVRSOTETKVEPFEVNPPLVIGSVGGLLLALITAAALYKLGFFKQYKD 1140
1141 MMSEGGPPGAEPQ 1153
1141 MMSEGGPPGAEPQ 1153

RESULT 4
AAU80252
ID AAU80252 standard; Protein; 1153 AA.
XX AAU80252;
XX AC AAU80252;
XX DT 15-JUL-2002 (first entry)
XX Human integrin 1 alpha-M subunit protein.
DE Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Misc-difference 499..500
FT /note= "Encoded by GGG CAG AGG"
XX WO200218583-A2.
XX 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US27227.
XX 01-SEP-2000; 2000US-229700P.
XX (BLOO-) CENT BLOOD RES INC.
XX Springer TA, Shimoaka M, Lu C;
PI

XX WPI, 2002-382964/41.
DR N-PSDB; ABK50046.
XX
PT Modified integrin-I or integrin I-like domain polypeptide useful as an
PT immunogen to produce antibodies specific to polypeptide, comprises a
PT disulfide bond such that polypeptide is stabilized in a desired
PT conformation -
XX
PS Disclosure; Page 109-112; 112pp; English.
XX
CC This invention relates to a modified integrin-I or integrin I-like
CC domain polypeptide comprising at least one disulfide bond so that the
CC domain is stabilized in a desired conformation. The polypeptide of
CC the invention may have antiinflammatory or immunosuppressive activities.
CC The polypeptides of the invention have an open conformation and are
CC useful as immunogens to produce antibodies that selectively bind to
CC integrin I-domain; and for identifying a modulator of integrin activity,
CC or of interaction of an integrin and a cognate ligand. The polypeptide
CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
CC is useful for treating or preventing an integrin mediated disorder which
CC is an inflammatory or autoimmune disorder in a subject and for
CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
CC disease, nephritis; human immunodeficiency virus (HIV), myocardial
CC infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis.
CC A therapeutic composition comprising the peptide of the invention is
CC useful for treating an integrin mediated disorder in a subject. The
CC polypeptides and/or active or antigenic fragments are useful as
CC reagents for diagnosis of integrin-mediated disorders. The present
CC sequence represents the human integrin-1 alpha-M protein subunit used to
CC generate the mutant polypeptides of the invention.
XX
SQ Sequence 1153 AA;
Query Match 100.0%; Score 5956; DB 23; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRLLLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAANQR 60
DB 1 MALRLLLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAANQR 60
QY 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVK 120
DB 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVK 120
QY 121 GLCFLGSNLROOPKFPALRGCPQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMBQL 180
DB 121 GLCFLGSNLROOPKFPALRGCPQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMBQL 180
QY 181 KKSRTLFSLMQYSEEFRIHFTFKEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KKSRTLFSLMQYSEEFRIHFTFKEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDGAFRSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDGAFRSEKSRQEL 300
QY 301 NTIASKPPRDHVQVNNFALAKTIONQUREKIFAIEGTQTGSSSFHEMSQEGFSAIT 360
DB 301 NTIASKPPRDHVQVNNFALAKTIONQUREKIFAIEGTQTGSSSFHEMSQEGFSAIT 360
QY 361 SNGPLLSTVGSYDAGGVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
DB 361 SNGPLLSTVGSYDAGGVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
QY 421 LGAPRYQHITGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHITGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
QY 481 HYYEOTRGQSVQVCPPLPRQARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEOTRGQSVQVCPPLPRQARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDVA 540

QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGGODLTMDGLV 600
DB 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGGODLTMDGLV 600
QY 601 DLTGAGQGHVLLLSRQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTGAGQGHVLLLSRQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAVFNETKSTRQTOVLGLTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAVFNETKSTRQTOVLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRVLAEDAQRFTALFPPEKNCNGNDNICODD 780
DB 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRVLAEDAQRFTALFPPEKNCNGNDNICODD 780
QY 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDCGDSYRQVTFPPFLDLSYRKVSTLQORS 840
DB 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDCGDSYRQVTFPPFLDLSYRKVSTLQORS 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNMPTNKTEFOLELPVKYAVVMVTSHGVTSTKYLNTFASENTSVMQHQYQVSN 960
DB 901 NVTSENNMPTNKTEFOLELPVKYAVVMVTSHGVTSTKYLNTFASENTSVMQHQYQVSN 960
QY 961 LGQRSPLISLVLVPLVRNQTIVWDRQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGQRSPLISLVLVPLVRNQTIVWDRQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCISIAVCQRIQCIPFPGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVF 1080
DB 1021 VNCISIAVCQRIQCIPFPGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVF 1080
QY 1081 TLLPQQAQFVRSQVTSQTKVEPEVNPPLPIVGVSSVGGLLLALITAALYKLGFPKQYKD 1140
DB 1081 TLLPQQAQFVRSQVTSQTKVEPEVNPPLPIVGVSSVGGLLLALITAALYKLGFPKQYKD 1140
QY 1141 MWSEGGPPGAPQ 1153
DB 1141 MWSEGGPPGAPQ 1153
RESULT 5
AAO14428
ID AAO14428 standard; protein; 1153 AA.
XX
AC AAO14428;
XX
DT 03-MAY-2002 (first entry)
XX
DE Integrin Mac-1 alpha subunit.
XX
KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.
XX
OS Unidentified.
XX
XX WO200204521-A2.
PN
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US21805.
XX
PR 07-JUL-2000; 2000US-216600P.
XX

PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX (BLOO-) CENT BLOOD RES.
PI Springer T;
XX WPI; 2002-148167/19.
XX
XX New integrin I domain protein having alteration in at least 2
PT noncontiguous regions and exits in an open conformation, useful for
PT treating, preventing or suppressing inflammatory or immunological
PT disorders
XX
XX Example 1; Fig 1F; 90pp; English.
XX
XX The invention comprises structurally biased variant integrin inserted (I)
CC domain proteins, wherein the alterations to the protein occur in at least
CC two noncontiguous regions. Specifically the variant integrin I domain
CC proteins are structurally biased to exist in the open conformation,
CC thereby altering the binding ability of the protein. The invention also
CC comprises nucleic acids encoding the variant integrin I domain proteins.
CC The integrin I domain proteins and nucleic acids are useful for treating,
CC preventing or suppressing integrin related inflammatory and immunological
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC proteins and nucleic acids can also be used for treating: ischaemia/
CC reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
CC infection; and cancer. The variant integrin I domain nucleic acids and
CC proteins may be used in gene therapy, as vaccines and to screen for
CC bioactive agents. The present amino acid sequence represents the Mac-1
CC alpha subunit of integrin.
XX
SQ Sequence 1153 AA;
Query Match 100.0%; Score 5956; DB 23; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRVLLLTALTCHGFNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQR 60
DB 1 MALRVLLLTALTCHGFNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQR 60
QY 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTSENTYVK 120
DB 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTSENTYVK 120
QY 121 GLCLFPGNLRRQPOKFPPEARLGCPOEDSDIAFLDGSIIIPHDPRMKFEVTVMEQL 180
DB 121 GLCLFPGNLRRQPOKFPPEARLGCPOEDSDIAFLDGSIIIPHDPRMKFEVTVMEQL 180
QY 181 KKSRTLFLSMQYSEBFRTHFTKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KKSRTLFLSMQYSEBFRTHFTKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFLVITDGEFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQEL 300
DB 241 ITNGARKNAFLVITDGEFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQEL 300
QY 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360
DB 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360
QY 361 SNGPLLTSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIIILNRVQSLV 420
DB 361 SNGPLLTSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIIILNRVQSLV 420
QY 421 LCAPRYQHILGVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LCAPRYQHILGVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
QY 481 HYYEQRGGQSVSCPLPRQGRARWOCDAVLVYGEQGPWRFGAALTIVLGDVNGDKLTDVA 540
DB 481 HYYEQRGGQSVSCPLPRQGRARWOCDAVLVYGEQGPWRFGAALTIVLGDVNGDKLTDVA 540
QY 541 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRAGSKLSPRLQYFQGSLSGGDLTMDGLV 600

DB 541 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRAGSKLSPRLQYFQGSLSGGDLTMDGLV 600
QY 601 DLTVGAQGHVLLRSQPVLRVKAINEFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTVGAQGHVLLRSQPVLRVKAINEFNPVARNVFECDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLRREGIQSVVYTDLALDSGRPHSRVFNETKNSRTRQTQVLGTQTCTETKLQLP 720
DB 661 STRDLRREGIQSVVYTDLALDSGRPHSRVFNETKNSRTRQTQVLGTQTCTETKLQLP 720
QY 721 NCIEDPVPSPVLRINFLNFSLVGTPLSAFGLRPLVLAEDAQRLLFTALPFFKNCNDNICQDD 780
DB 721 NCIEDPVPSPVLRINFLNFSLVGTPLSAFGLRPLVLAEDAQRLLFTALPFFKNCNDNICQDD 780
QY 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPPFDLSYRKVSTLQNORS 840
DB 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPPFDLSYRKVSTLQNORS 840
QY 841 QRSWRLACESASSTEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWRLACESASSTEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENMPRTNKTEFOLELPVKYAYVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
DB 901 NVTSENMPRTNKTEFOLELPVKYAYVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
QY 961 LGORSUPISLVFLVPVRLNQTWIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGORSUPISLVFLVPVRLNQTWIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCISIAVCQRIQCDIPFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVF 1080
DB 1021 VNCISIAVCQRIQCDIPFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVF 1080
QY 1081 TLLPGQGAFAVRSQVSTETKVPFEPVNPPLIVGSSVGLLALLAIIITAAALYKLGFFKQYKD 1140
DB 1081 TLLPGQGAFAVRSQVSTETKVPFEPVNPPLIVGSSVGLLALLAIIITAAALYKLGFFKQYKD 1140
QY 1141 MMSEGGPPGPAEPQ 1153
DB 1141 MMSEGGPPGPAEPQ 1153
RESULT 6
AAR04136
ID AAR04136 standard; protein; 1153 AA.
XX AAR04136;
XX AC
XX 25-MAR-2003 (updated)
DT 07-SEP-1990 (first entry)
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX DE
XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX KW non-specific defence system; integrin gene superfamily.
XX non-synthetic.
XX OS
XX Key Location/Qualifiers
FH Modified-site 86..88
FT /*label= putative N-glycosylation site
FT Modified-site 240..242
FT /*label= putative N-glycosylation site
FT Modified-site 391..393
FT /*label= putative N-glycosylation site
FT Modified-site 469..471
FT /*label= putative N-glycosylation site
FT Modified-site 693..695
FT /*label= putative N-glycosylation site
FT Modified-site 697..699
FT /*label= putative N-glycosylation site

FT Modified-site 735..737 /*label= putative N-glycosylation site
 FT Modified-site 802..804 /*label= putative N-glycosylation site
 FT Modified-site 881..883 /*label= putative N-glycosylation site
 FT Modified-site 901..903 /*label= putative N-glycosylation site
 FT Modified-site 912..914 /*label= putative N-glycosylation site
 FT Modified-site 941..943 /*label= putative N-glycosylation site
 FT Modified-site 947..949 /*label= putative N-glycosylation site
 FT Modified-site 979..981 /*label= putative N-glycosylation site
 FT Modified-site 994..996 /*label= putative N-glycosylation site
 FT Modified-site 1022..1024 /*label= putative N-glycosylation site
 FT Modified-site 1045..1047 /*label= putative N-glycosylation site
 FT Modified-site 1051..1053 /*label= putative N-glycosylation site
 FT Modified-site 1076..1078 /*label= putative N-glycosylation site
 FT region 1..16 /*label= signal peptide
 FT region 1106..1134 /*label= putative transmembrane region
 XX
 PN EP364690-A.
 XX
 PD 25-APR-1990.
 XX
 PF 17-AUG-1989; 89EP-0115159.
 XX
 PR 23-AUG-1988; 88US-0235353;
 PR 09-MAR-1989; 89US-0321239.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Springer TA, Corbi A;
 DR WPI; 1990-125938/17.
 DR N-PSDB; AAQ04043.
 XX
 XX New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
 PT inflammation and viral infections, and in diagnosis
 XX
 PS Disclosure; Page ?; ?pp; English.
 XX
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue.
 CC Mac-1 is a member of the Integrin Gene superfamily.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1153 AA;
 Query Match 99.8%; Score 5946; DB 11; Length 1153;
 Best Local Similarity 99.9%;
 Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MALRVLLTALTLCGFLNLTENAMTFQENARGFQGVVLOGSRVVVVGAPQEIIVAANQR 60
 DB 1 MALRVLLTALTLCGFLNLTENAMTFQENARGFQGVVLOGSRVVVVGAPQEIIVAANQR 60
 QY 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
 DB 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120

QY 121 GLCFLFGSNLRQQPKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
 DB 121 GLCFLFGSNLRQQPKFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
 QY 181 KSKTSLSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
 DB 181 KSKTSLSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
 QY 241 ITNGARKNAFKILVITDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSKQEL 300
 DB 241 ITNGARKNAFKILVITDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSKQEL 300
 QY 301 NTIASKPPDRHVFQVNNFEALKTIONQRLREKIPIAIEGTQTGSSSSFEHMSQEGFSAIT 360
 DB 301 NTIASKPPDRHVFQVNNFEALKTIONQRLREKIPIAIEGTQTGSSSSFEHMSQEGFSAIT 360
 QY 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTFRVDSMDNDAYLGYAAAIIILNRVOSLV 420
 DB 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTFRVDSMDNDAYLGYAAAIIILNRVOSLV 420
 QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAP 480
 DB 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAP 480
 QY 481 HYTEQTRGGQVSVCPFRGQARWQCDAVLYGEGQPMGRFGAALTIVLGDVNGDKLTDVA 540
 DB 481 HYTEQTRGGQVSVCPFRGQARWQCDAVLYGEGQPMGRFGAALTIVLGDVNGDKLTDVA 540
 QY 541 IGAPGEDNDRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDUTMDGLV 600
 DB 541 IGAPGEDNDRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDUTMDGLV 600
 QY 601 DLTVGAQGHVLLLRSPQVLRVKAIMEFNPREVAARVFNCECDVVKGEAGVRVCLHVQK 660
 DB 601 DLTVGAQGHVLLLRSPQVLRVKAIMEFNPREVAARVFNCECDVVKGEAGVRVCLHVQK 660
 QY 661 STRDLREGQIQSVVITYDLALDSGRPHSRVFNFTKSTRQTOVLGLTQTCETLKLQLP 720
 DB 661 STRDLREGQIQSVVITYDLALDSGRPHSRVFNFTKSTRQTOVLGLTQTCETLKLQLP 720
 QY 721 NCIEDPVPSPVLRNLFSLVGTPLSAFGLNRPVLAEDAQRPLTALFPPEKNGCNDNI CDD 780
 DB 721 NCIEDPVPSPVLRNLFSLVGTPLSAFGLNRPVLAEDAQRPLTALFPPEKNGCNDNI CDD 780
 QY 781 LSITFSEMSLDCLVVGPPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQORS 840
 DB 781 LSITFSEMSLDCLVVGPPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQORS 840
 QY 841 QRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
 DB 841 QRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
 QY 901 NVTSENMPRTNKTFOLELPVKYAVVMVTSHGVSSTKYLNTASNTSRVMQHOYQVSN 960
 DB 901 NVTSENMPRTNKTFOLELPVKYAVVMVTSHGVSSTKYLNTASNTSRVMQHOYQVSN 960
 QY 961 LGQSLPLISLVFLVPRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
 DB 961 LGQSLPLISLVFLVPRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
 QY 1021 VNCISIAVCQRIQCDIPFFGQIEEFNATLKGNLSDFWIKTSHNHLIIVSTAEIIFNDSVF 1080
 DB 1021 VNCISIAVCQRIQCDIPFFGQIEEFNATLKGNLSDFWIKTSHNHLIIVSTAEIIFNDSVF 1080
 QY 1081 TLLPGQAFVRSQETETKVEPEVFNPLPIVGVSSVGGLLLLALITAALYKLGFKKQYKD 1140
 DB 1081 TLLPGQAFVRSQETETKVEPEVFNPLPIVGVSSVGGLLLLALITAALYKLGFKKQYKD 1140
 QY 1141 MMSEGGPPGAEPO 1153
 DB 1141 MMSEGGPPGAEPO 1153

Db 903 ENNTPTSKTTFQLELPKYAVYTVSSHEQFTKYNFSESEKESHVAMHYQVNNLQ 962
 Qy 964 RSLPTSLVLPVRLNQTIVDRPOVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNC 1023
 Db 963 RDLPSVINFVPELVNQAAMDVEVSHPOQPSLRCSQKTIAPASDPLAHIQKNPVLC 1022
 Qy 1024 SIACVORIQCDDIPFGIQEENATLKGNSLFDWYIKTSHNHLIVSTAELFNDVSFTLL 1083
 Db 1023 SIAGCLRPDCVPSPSQEELDFTLKGNSLFGWRQILQKKVSVVVAEITFDTSVYSQL 1082
 Qy 1084 PGQAFVRSQETKVEPEVNPPLIVGSSVGGLLILALITAAIKYKGFPPKRYKQDMS 1143
 Db 1083 PGOERFRAQTTTLEKYKVHNPPLIVGSSIGALLILALITAVLYKVGFPPKRYKENME 1142
 Qy 1144 E 1144
 Db 1143 E 1143

RESULT 8

AAW65091
 ID AAW65091 standard; Protein; 1163 AA.

XX AC AAW65091;

XX DT 28-SEP-1998 (first entry)

XX DE Human Beta-integrin CD11c subunit protein.

XX KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
 lung inflammation; acute respiratory distress syndrome; CD11c subunit;
 rheumatoid arthritis.

XX OS Homo sapiens.

XX PN US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-0485618.

XX PR 07-JUN-1995; 95US-0485618.

XX PR 23-DEC-1993; 93US-0172497.

XX PR 05-AUG-1994; 94US-0286889.

XX PR 21-DEC-1994; 94US-0362652.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin WM, Van DER VIEREN M;

XX DR WPI; 1998-206565/18.

XX PT Screening assay for modulators of integrin binding - using

XX PT immobilised or labelled alpha-d polypeptide, useful for, e.g.

XX PT treating type-I diabetes

XX PS Example 5; Fig 1A-D; 106pp; English.

XX CC This sequence represents a human beta-integrin CD11c subunit which is
 used to describe a method for identifying compounds that modulate the
 interaction of the beta-integrin alpha-d subunit with a binding partner
 of alpha-d which involves contacting an alpha-d polypeptide with an
 alpha-d binding partner, one of which is immobilised and the other of
 which is labelled, in the presence of a test compound, and determining if
 the compound affects binding between the alpha-d polypeptide and alpha-d
 binding partner, where the alpha-d polypeptide is alpha-d or its fragment
 comprising the cytoplasmic, transmembrane or extracellular domain of
 alpha-d. Compounds that modulate alpha-d binding could be used to treat
 diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
 asthma, psoriasis, lung inflammation, acute respiratory distress
 syndrome and rheumatoid arthritis.

XX SQ Sequence 1163 AA;
 Query Match 58.4%; Score 3480; DB 19; Length 1163;
 Best Local Similarity 61.4%; Pred. No. 7.5e-281;
 Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;
 Qy 5 VLLLTALTLCGHNLTENAMTFQENARGFQSVVQLQGSVVVVGAPQEIIVAAQNGSLY 64
 Db 8 LLLFTALATSLGHNLTETELTAFRVDSAGFGSDVQVYANSVVVGAPQKIIAANQIGLY 67
 Qy 65 QDYSTGSCSPIKLVQVPEAVNMSLGLSLAATSPQLLACGPTVHOTCSENTYVYKGLCF 124
 Db 68 QCGYSTGACSPICLVQVPEAVNMSLGLSLAATSPQLLACGPTVHOTCSENTYVYKGLCF 127
 Qy 125 LFGSNLRQOPQKPEALRGCPQSDSDIAFLIDCGSGIIPHDPRMKEFVSTVMEQLKSK 184
 Db 128 LLGPT--QLTQRLPVSRQECPEQDIFVLIDGSGSISSRNFMFMFVRAVISQFORPS 185
 Qy 185 TLFSLMQYSEFRIHFTFEFONNPNRSLVKPITOLLGRTHATGIRKVVRELFNITNG 244
 Db 186 TQFSLMQFNSKFTHTFEFRTSNPLSLASVHQLQGFYTATATQNVVHRLFHASYG 245
 Qy 245 ARKNAFKILVITDGEKFGDPLGYDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIA 304
 Db 246 ARDAIKILIVITDGKKEGSDLDYKDVIPMADAAGIIRYAIYGLAFQNRNSMKELNDIA 305
 Qy 305 SKPRDRHVQVNNFEALKTIONQLREKIPIAIEGTOTGSSSSSFEHMSQEGFSAITNGP 364
 Db 306 SKPSQEHIFKVEDFDALQIQNLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTDPGP 365
 Qy 365 LLGTGSGYDWAGVFLYTSKSTFTNMTRVDSMDNDAYLGAAAIILNRVQSLVIGAP 424
 Db 366 VLGAVSFTWSGAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELAKWGVQSLVIGAP 425
 Qy 425 RYQHIGLVAMFRQNTGWMESNANVKGITQIGAYFGASLCSDVDVNSGSTDLVLGAPHYE 484
 Db 426 RYQHIGKAVIFIQVSRQWRMAEVIQTQIGSYFGASLCSDVDVDTGSDTLVLGAPHYE 485
 Qy 485 QTRGGQVSVCPPLRGQARQWQCDVLYGEOGQWGRFGAALTVLGVGVNGDKLTDVAIGAP 544
 Db 486 QTRGGQVSVCPPLRGWR-RWVCDVLYGEOGQWGRFGAALTVLGVGVNGDKLTDVIGAP 544
 Qy 545 GEEDNRGAVLPHGTSGSGISPSHSORIASGKLSPLQYFGOSLSCGODLTGMLVDLTV 604
 Db 545 GEBENRGAVLPHGVLPISPSHSORIASGKLSPLQYFGOSLSCGODLTGMLVDLTV 604
 Qy 605 GAQGHVLLRSQPLRVKAIMBENPREVARNVPECNDQVVKGEAGEVRVCLHVOKSTRD 664
 Db 605 GARGQVLLLRTPVLVWGVSMQFIPAEIPRSAFECEQVSEQTLVQSNTCLYIDKRSKN 664
 Qy 665 RLREGQISVVTYDLALDSCRPHSRVFNETKNSRRTQTVLGTQTCETLKLQLPNCIE 724
 Db 665 LLGSRDLQSSVTLDLALAPGLSPRAIFQETKNSLSRVVRLGLKAHCENFNLLPSCVE 724
 Qy 725 DPVSPVILNPLNSLVTPLSAFQNLAPVLAEDAQRLFTALPPEKNCNNDNICODDISIT 784
 Db 725 DSVIPITILNFTLVGKPLLAFLNPLAALAAQRYFTASLPPEKNCNNDNICODNIS 784
 Qy 785 FSPMSLDCLVGGPREFNVTVTVRNDEDSYRTQVTFPPPLDLSYRKVSTLQVRSORSW 844
 Db 785 FSPFGLKSLLVGNSLNAEVMVWVNDGDSYGTITFSPAGLSYRYVAGQKQQRSL 844
 Qy 845 RLACESASSTEVSGALKSTSCSINHPIFPIENSEVTNITFDVDSKASLGNKLLKANVTS 904
 Db 845 HLTC--CSAPVSGQGTWSTSCRINHIFRGAQITFLATFDVSPKAVGLDRLLLIANVSS 902
 Qy 905 ENNMPTNKTEFOLELPVKVAVYVMTSHCVSTKYLNFTAS-ENTSRVMOHQVQVSNLQ 963
 Db 903 ENNIPRTSKTIFQLELPVKVAVYVMTSHCVSTKYLNFTAS-ENTSRVMOHQVQVSNLQ 962
 Qy 964 RSLPISLVFLVRLNQTIVDRPOVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNC 1023

Db 963 RDLPSINFWVPVELNQEAVMDVEVSHFQNPDLRCSSEKIAPPASDFLAHIQKNPVLDC 1022
QY 1024 STAVCORIQCDIPFFGIQEFNATLKGNSLFDWYIKTSHNHLIIYSTAILFENDSVFTLL 1083
Db 1023 SIAGCLRFCDVPFSFVSQBELDFTLKGNSLFGWVROILQKKVSVSVSVAIIIFDTYSISOL 1082
QY 1084 PGGAFAVRSGOTKVPFFVNPPLPIVGVSSVGGILLALLALITAAALYKLGFFKQYKQWMS 1143
Db 1083 PGGAFAVRSGOTKVPFFVNPPLPIVGVSSVGGILLALLALITAVLYKVGFKQYKQWME 1142
QY 1144 E 1144
Db 1143 E 1143

RESULT 9
AAB07361.
ID AAB07361 standard; Protein; 1163 AA.
AC AAB07361;
XX
DT 17-JAN-2001 (first entry)
DE Human CD11c protein sequence.
KW Human; macrophage infiltration inhibition; alpha d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX
OS Homo sapiens.
XX WO200029446-A1.
PN
XX 25-MAY-2000.
XX 16-NOV-1999; 99WO-US27139.
XX 16-NOV-1998; 98US-0193043.
PR 08-JUL-1999; 99US-0350259.
XX (ICOS-) ICOS CORP.
PA Gallatin MW, Van Der Vieren M;
PI WPI; 2000-387751/33.
XX
XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous
PT system injury sites -
XX
XX Example 5; Fig 1; 270pp; English.
XX
XX Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins. Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11c. This
CC sequence was used in an alignment to identify a novel beta2 integrin
CC alpha subunit: alpha d (AA060014 and AAB07359). The present sequence has
CC approximately 66% identity to the protein sequence of alpha d. The
CC Alpha d gene and protein may be useful in therapy for diseases linked
CC to aberrant alpha d function e.g. Type I diabetes, atherosclerosis,
CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
CC adhesion deficiency (LAD). In addition, anti-alpha d monoclonal
CC antibodies may be used in the inhibition of macrophage infiltration at
CC the site of a central nervous system injury. The monoclonal antibodies
CC can also be used to detect and diagnose Crohn's disease.

XX Sequence 1163 AA;
QY Query Match 58.4%; Score 3480; DB 21; Length 1163;
Best Local Similarity 61.1%; Pred. No. 7.5e-281;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;
Db 5 VLLTALTLCHGNLDTENAMTFQENARGFGQSVVQOGSRVVGVGAPQEIIVAAHQSGSLY 64
8 LLLFTALATSLGFLNLTDELTAFRVDSAGFGDSVVQYANSVVVGAPQKIIAANQIQGLY 67
QY 65 QCDYSTGSCBPIRLOVPVEAVNMSLGLSLAATTSPOLLACGPTVHQCSTNTYVKGLCF 124
68 QCGYSTGACEPIGLQVPEAVNMSLGLSLASTTSPQLACGPTVHHCGRNMYLTGUCF 127
QY 125 LFGSNLRQOPKFPALRGCCPOEDSDIAFLIDGSGSIIPHDPRRMEKEFVSTVMEOLKSK 184
128 LLGPT--QLTQLPVSROECPRQEQDIVFLIDGSGSISRNFATMNFVRAVISQFQPS 185
QY 185 TLFLSMQVSEEFRIHFTFKBFONNPNRSLVKPIITQLLGRTHATGIRKVVRELNYING 244
186 TQFSLMQFSNKFQTHFTFEFRRTSNPLSLASVHQLQGFYTTATAIQNVHRLPHASYG 245
QY 245 ARKNAPKILVITDCEKFGDPLGVEDVIPADREGVIRVIGVGDAFRSEKROELNTIA 304
246 ARDAIKILIVITDCKKGGDSLIDYKDVIPMAAAGIIRIAYGLVAFQNRNWKELNDIA 305
QY 305 SKPPRDHVPQVNNFEALKTIONQLREKIPAIETGOTGSSSSFEHEMSQEGFAAITSNGP 364
306 SPSQSEHFKVEDFALKDIONLKEKIPAIETGOTGSSSSFEHEMSQEGFAAITSNGP 365
QY 365 LLSTVGSVDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVGLAP 424
366 VLGAVGSFTWGGAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVSQSLVGLAP 425
QY 425 RYQHICLVAMPFRQNTGMWESNANVKGTOIGVFGASLCSVDVDSNGSTDLVLIGAPHYYE 484
426 RYQHIGKAVIFTQVSRQWRMKAIEVIGTQIGSFGASLCSVDVDTGSDTLVLIGAPHYYE 485
QY 485 QTRGGQVSVCPPLRGORARWQCDVLYGEGQGPWGRFGAALTIVLGVNGDKLTDVAIGAP 544
486 QTRGGQVSVCPPLRGWR--RWCDVLYGEGQHPWGRFGAALTIVLGVNGDKLTDVVGAP 544
QY 545 GEDNRGAVYLFHGTSGSGISPSHSORIASGKLSPLQYFGQSLSGQDLTMDGLVDLTV 604
545 GEBENRGAVYLFHGTSGSGISPSHSORIASGKLSPLQYFGQSLSGQDLTMDGLVDLTV 604
QY 605 GAQGHVLLRSQPLRVKAIMENPREVARNVFECDNVVYKGEAGEVRVCLHVOKSTRD 664
605 GARGQVLLLRTRPVLWGVSMQFIPAEIPRSAFECEQVQVSEQTLVQSNICLYIDKRSKN 664
QY 665 RLREGQIQSVVYDLDALDSGRPHSRVAFNETKNSTRQVQLGLTQTCETLKLQLPNCIE 724
665 LLGSRDLQSSVTLDLALAPGLSPRAIFQETKRSLSRVRLVGLKAHCENFLLPSCVE 724
QY 725 DPVSPVILRLNFSVLGTPLSAFNLNRPVLAEDAQRLFTALFPPEKNGCNDNI CODDLSIT 784
725 DSVIPIILRLNFTLVGKPLLAFLNRLPMLAALAAQRYFTASLPPEKNGCADHICODNLGIS 784
QY 785 FSPMSLDCLVVGPREFNTVTVRNDEGDSYRQVTFPFFPLDLISYKRVSTLQNRORSW 844
785 FSPFGLKSLVGNLELNAEVMWMDGSDSYGTTITFSPHAGUSYRVASGQKQKQURSL 844
QY 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
845 HLTC--CSAPVSGQGTWSTSCRINHPIFRGAQITFLATFDVSPKAVGLDRLLLIANVSS 902
QY 905 ENNMPRTNKTFFOLELVPKVAVYVMTVSHGVSTKYLNFPTAS--ENTSRMHOHQVSNLQ 963
903 ENNIPRTSKTIFOLELVPKVAVYVMTVSHGVSTKYLNFPTAS--ENTSRMHOHQVSNLQ 962
QY 964 RSLPISLFLVPLVPLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNC 1023

Db 963 ROLPVSINFWVPELVNQEAVMDVESHVPHQNSLRCSSEKIPAPPASDFLAHQKPNVLDL 1022
QY 1024 SIACVQRIQDIPFFQIOBEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSTLL 1083
Db 1023 SIAGCLRPFCVPSFQSELDFTLKNLSFGWVRQILQKKVSVVSAEIIIDTSVYSOL 1082
QY 1084 PQGAPVRQSTETKVEPFVNPPLPIUVSSVGGLLLLALITAAALYKLGFFRKQYKDMMS 1143
Db 1083 PQGEAFMRAQTITVLEKYVHNPIPLIVGSSIGGLLLALITAVLYKVGFRRQYKEMME 1142
QY 1144 E 1144
Db 1143 E 1143

RESULT 10
ID ABG61470 standard; Protein; 1163 AA.
XX ABG61470;
XX 27-AUG-2002 (first entry)
DT Human Beta2 integrin alphaCD11c subunit.
DE
XX Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammation;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX Homo sapiens.
OS
XX WO200230980-A2.
PN
XX 18-APR-2002.
PD
XX 15-OCT-2001; 2001WO-US32059.
PF
XX 13-OCT-2000; 2000US-0688307.
PR
XX (ICOS-) ICOS CORP.
PA
XX Gallatin WM, Van Der Vieren M;
PI
XX WPI; 2002-463260/49.
DR
XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -
XX
PS Example 5; Page 194-198; 270pp; English.
XX
XX The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an
CC anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC spinal cord injury victim. The method also involves the use of
CC a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion
CC molecule, vascular cell adhesion molecule). The method is useful for
CC promoting locomotor recovery, inhibiting locomotor damage, limiting
CC locomotor impairment, or limiting autonomic and sensory dysfunction
CC following spinal cord injury. In particular, the spinal cord injury
CC comprises compression of the spinal cord. The antibodies are also useful
CC for reducing inflammation at the site of a central nervous system injury.
CC The specification also details the identification of Beta2 integrin
CC alphaD cDNAs and proteins, for use in raising the antibodies. Beta2
CC integrins are implicated in diseases such as LAD (leukocyte adhesion
CC deficiency, inflammatory response, diabetes, multiple sclerosis,

CC arthritis, graft atherosclerosis, inflammatory bowel disease,
CC Crohn's disease, ulcerative colitis, immune complex alveolitis,
CC and leukaemia. The present sequence is a Beta2 integrin alpha subunit
CC sequence included for comparison with the Beta2 integrin alphaD protein
CC sequences.
XX
SQ Sequence 1163 AA;
Query Match 58.4%; Score 3480; DB 23; Length 1163;
Best Local Similarity 61.1%; Pred. No. 7.5e-281;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;
QY 5 VLLTALTCHGFNLDENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAAQNGSLY 64
Db 8 LLLFTALATSLGFNLDTEELTAFRVDSAGFGDSVVQYANVWVGAPQKIIAANGIGLY 67
QY 65 QCDYSTGCEPIRLQVPVEAVNMISGLSLAATTSPOLLACGTPVHQCSTENTYKGLCF 124
Db 68 QCGYSTGACEPIGLQVPPPEAVNMISGLSLASTSPSLLACGTPVHHECGRNMVLTGJCF 127
QY 125 LFGSNLRQOPQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOLKSK 184
Db 128 LLGPT--QLTORLPVSRQECPRQEQDIVFLIDGSGSISSRFATMNFVRAVISQFQPS 185
QY 185 TLFLSMQYSEFRITHFTKFEQNNPNRSLVKPIQLLGRTHATGIRKVVRELFNITNG 244
Db 186 TQFSLMQPSNKFQTHFTFEFRRTSNPLSLASVHQLQGFVTTATAIQNVVHRLPHASVG 245
QY 245 ARKNAFKLVITDGEKCDPLGYEDVITPEADREGVIRYVIGVGDAFSEKSRQELNIA 304
Db 246 ARDDAIKILVITDGEKCDPLGYEDVITPEADREGVIRYVIGVGDAFSEKSRQELNIA 305
QY 305 SKPPRDHVFQVNNFEALKTIONLREKIFAIEGTOTGSSSFEHEMSOEGFSAAITNCP 364
Db 306 SKPSQEHFKVEDFDALKDIONLREKIFAIEGTETISSSFEHEMSOEGFSAAITNCP 365
QY 365 LLSTVSGYDWAGGVLYTSKEKSTFINMTRVDSNDMDAYLGYAAIILNRNVQSLVLGAP 424
Db 366 VLGA VGSFTWSGGAFLYPPNMSPTFINMSQENVMRDSYLGYSTELALWKGVQSLVLGAP 425
QY 425 RYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYIE 484
Db 426 RYQHIGKAVIFIQVSRQWRMKAENVIGTQIGSYFGASLCSVDVDTGSDTDLVIGAPHYIE 485
QY 485 QTRGGQSVVCPPLRGORARWOCDAVLYGEGQGPWRFGAALTVLGDVNGDKLTDVAIGAP 544
Db 486 QTRGGQSVVCPPLRGWR--RWCDAVLYGEGQGPWRFGAALTVLGDVNGDKLTDVIGAP 544
QY 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLV 604
Db 545 GEEENRGAVYLFHGVLPSPISPSHSQRIAGSKLSPRLQYFGQALSGGQDLTQDGLVDLAV 604
QY 605 GAGQHVLLLRQOPVLRVKAIWEPNPREVARNVFECDQVVKGEAGEVRLVHVKSTRD 664
Db 605 GARGQVLLLRTRPVLWVGVSQGFPAEIPRSAFECEQVWSQTLVQSNICILYDKRSN 664
QY 665 RLREGQIOSVVTYDLALDSGRPHSRVFNENKSTRTQTVLGLTQTCETLKLQPNCTE 724
Db 665 LIGSRDQSSVTDLALAPGRISPRAI FOETKNRSLSRVRVIGLKAHCENFLLPSCVE 724
QY 725 DPVSPVILRLNFSVLGTPLSAFNGRLPVLAEQAQLFTALFPFKNCGNDNICQDLSIT 784
Db 725 DSVIPIILRLNFTLVGKPLLAFLNLRPLAALAAQRYFTASLPFKNCGADHICQDNLGIS 784
QY 785 PSFMSLDCVLVGGPREFNVTVTVRNDGEDSVRTQVTFPPDLDSYKRVKSTLQNRQSRW 844
Db 785 PSFPGKSLILVNSLLENAEVMVMWMDGDSYCTTITFSHPAGLSYRYAEGQKQGLRSL 844
QY 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
Db 845 HLTC--CSAPVGSQGTWSTSCRINHLIFRGQAQITFLATFDVSPKAVGLDRLLIANVSS 902
QY 905 ENNMPRTNKTQLELPVKYAVYVMVTVSHGVSTKYLNTFAS--ENTSRVMQHQYQVSNLQ 963

Db 903 ENNIPRTSKTIFOLELVPKAVYIVSSHEQFTKYNLFSESEKESHVAMHYQVNNLQ 962
Qy 964 RSLPISLVFLVPRLNQTIVMDRPOVTFSENLSTCHTKERLPSHSDFLAELKAPVNC 1023
Db 963 RDLFVSINFVPVELNQEAVMDEVEVSHQNPQLSRCSSEKIAPPASDFLAHQNPVLD 1022
Qy 1024 SIACVQRIQCDIPFGQEEFNATLKNLGFDFWYIKTSHNHLIVSTAELFNDVFTLL 1083
Db 1023 SIACVQRIQCDIPFGQEEFNATLKNLGFDFWYIKTSHNHLIVSTAELFNDVFTLL 1082
Qy 1084 PQGAFVRSQTTETKVEFVNPPLIVGSGVGLLALLIATLALYKGLFFKQYKDMMS 1143
Db 1083 PQGAFVRSQTTETKVEFVNPPLIVGSGVGLLALLIATLALYKGLFFKQYKDMMS 1142
Qy 1144 E 1144
Db 1143 E 1143

RESULT 11

ABU07406
ID ABU07406 standard; Protein; 1163 AA.

XX AC ABU07406;
XX DT 28-JAN-2003 (first entry)

XX DE Protein differentially regulated in prostate cancer #9.

XX KW Prostate cancer; gene expression; differential regulation;
XX KW molecular marker; drug target; cancer detection; cancer diagnosis;
XX KW cancer staging; cancer grading; cancer assessing; cancer monitoring.

XX OS Homo sapiens.

XX PN WO200281638-A2.

XX PD 17-OCT-2002.

XX PF 08-APR-2002; 2002MO-US10824.

XX PR 06-APR-2001; 2001US-281731P.

XX PR 06-APR-2001; 2001US-281732P.

XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX PI Sun Z, Jay G;

XX PS WPI; 2003-058520/05.

XX PT Novel genes which are differentially regulated in prostate cancer,
XX PT useful for diagnosing prostate cancer in prostate tissue sample and
XX PT assessing therapeutic or preventive intervention in prostate cancer
XX PT patients -

XX PS Claim 1; Page 225-228; 416pp; English.

XX CC The invention describes genes (I) which are differentially regulated in
XX CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
XX CC sample comprising prostate tissue, which involves determining the number
XX CC of target genes which are differentially-regulated in the sample, where
XX CC the number is indicative of the probability that the sample comprises
XX CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
XX CC intervention in a subject having a prostate cancer, which involves
XX CC determining the expression levels in a sample comprising prostate tissue
XX CC of target genes which are differentially-regulated in prostate cancer.
XX CC Preferably, the expression levels of at least 10 genes are determined.
XX CC (I) is also useful for identifying agents that modulate a biological
XX CC activity of a polypeptide differentially-regulated in prostate cancer
XX CC cells, which involves contacting a polypeptide differentially-regulated
XX CC in prostate cancer cells with a test agent under conditions effective for
XX CC the test agent to modulate a biological activity of the polypeptide, and

CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.

XX Sequence 1163 AA;

Query Match 58.4%; Score 3478; DB 24; Length 1163;

Best Local Similarity 61.0%; Pred. No. 1.1e-280;

Matches 696; Conservative 136; Mismatches 303; Indels 6; Gaps 4;

Qy 5 VLLLTALTLCCHGNLDENAMTFQENARGFQGSVQLOQGSRVVVGAPQETVAANQSGSLY 64

Db 8 LLLFTALATSLGNLDTEELTAFRVDSAGFSDSVQVANSVWVVGAPQKITAANQSGGLY 67

Qy 65 QCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVKGLCF 124

Db 68 QCGYSTGACBPIGLQVPPPEAVNMSLGLSLASTTSPSOLLACGPTVHHECRNMYLTGLCF 127

Qy 125 LFGSNLRQOPKEPEALRGCPQEDSDIAFLIDCGSGTTPHDFRMKEFVSTVMEQLKSK 184

Db 128 LLGPT--QLTORLPVSRQCEPQEQDIVLIDGSGTSSSRNFATMNFVRAVISQFORPS 185

Qy 185 TLPSLMQYSEEFRIHFTFEFQNNPNRSLVKITQLLGRTHATATGIRKVVRELNFINTG 244

Db 186 TQSLMQFSNKFTHLTFEERTSNPLSLASVHQLQGYTATATQNVVHRLPHASYG 245

Qy 245 ARKNAPKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIA 304

Db 246 ARRDATKILIVITDGKKEGDTLDYKQVIMPADAAGIIRYAIGVGLAFQNRNSKELNDIA 305

Qy 305 SKPRDRHVQVNNFEALKTIONQLREKIFAIEGTQGTGSSSSFEHMSQEGFSAITSNP 364

Db 306 SKPSQEHIFKVEDFDALKDIQTQUREKIFPIEGTETSSSSFELEMAQEGFSVFTPDGP 365

Qy 365 LLSTVGSYDWAGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAP 424

Db 366 VLGNVGSFTWGGCAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALMKGVSQSLVGLAP 425

Qy 425 RYQHIGLVAMFRONTGHWESNANVKGTOIGAYFGAGSLCSVDVDSNGSTDLVLGAPHYE 484

Db 426 RYQHTGKAVIFTQVSRQWRMKAEBVTGTQIGSYFSPSLCSVDVDSGSTDVLVIGPPHYE 485

Qy 485 QTRGQVSVCLPRGQBARWQCDVLYGEOGPMGRFGAALTVLGDVNGDKLTQVATGAP 544

Db 486 QTRGAQVSVCLPRGQWR--RWMCDAVLYGEOGHPMGRFGAALTVLGDVNGDKLTQVATGAP 544

Qy 545 GEENRGAVALFCHTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTV 604

Db 545 GEENRGAVALFCHTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTV 604

Qy 605 GAQGHVLLRSQPVLRVKAIMEFNPRVARNVFCNDQVVKGEAGEVVRVCLHVQKSTRD 664

Db 605 GARGQVLLLRTPVLVWGVSMQFIPABIPRSAFECEQWSEQTLVQSNICLYIDKRSKN 664

Qy 665 RLREGQVSVTVYDLALDSCPHSRVAFNETKNSRQTVQLGTQTCETLKLQLPNCIE 724

Db 665 LLGSRDLQSSVTLDLALDPGRSLSPRATFQETKNRSLGRVRLGLKAHCENFNLPLPSGCV 724

```

QY 725 DPVSPVILNLSVLTPLSAGNLRPVLAEDAQRLLFTALFPFKNGCNDNIQODLSIT 784
Db 725 DSVTPITLNLFTLVGKPLAFNLRPLAADAQRYFTASLPFKNGGADHICQDNLGIS 784
QY 785 F5FMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPLDLSYRKVSTLQNSORSW 844
Db 785 F5FPGUKSLVGNLENAEVMVMNDGDSYGTITTFSPAGLSYRYVAEGQKQSL 844
QY 845 RLACESASSTEVSGALKSTSCSINHPIFFPENSEVTENITFDVDSKASLGNKLLKANVTS 904
Db 845 HLTCDAPVG--SQGTWTSISCRINHILIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSS 902
QY 905 ENNMPTNTKTFQLELPVKYAVVMVTVSHGVSTKYLNFAS-ENTSRVMQHOYQVSNLQ 963
Db 903 ENNMPTNTKTFQLELPVKYAVVMVTVSHGVSTKYLNFAS-ENTSRVMQHOYQVSNLQ 962
QY 964 RSLPISLVLVPVRLNOTVWDRPQVTFSENLSSTCHTKERLPSPHSDFLAELRKAPVNC 1023
Db 963 RDLPSVSNFVPELVNQEAVMMDVEVSLPQNPSLRCSSEKIAGPASDFLAHTQKNPVLC 1022
QY 1024 STAVCORIOCDIPFFGIGQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLL 1083
Db 1023 STAGCLRFCDVPFSVQSELDFTLKNLSFGWROILQKKVSVSVAIITDTSVYSOL 1082
QY 1084 PGQAFVRSQETKVPFFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKXMS 1143
Db 1083 PGQAFVRSQETKVPFFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKXMS 1142
QY 1144 E 1144
Db 1143 E 1143

RESULT 12
AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX AC AAR78166;
XX DT 28-DEC-1995 (first entry)
XX DE Human beta-2 integrin alpha-d.
XX KW Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis;
XX KW inflammatory bowel disease; asthma.
XX OS Homo sapiens.
XX FH Key
XX FT Location/Qualifiers
XX FT Domain
XX FT 17..1108
XX FT /note= "extracellular domain"
XX FT Region
XX FT 150..352
XX FT /note= "this region is homologous to the insert
XX FT common to Chila,b,c and may be a site for
XX FT interaction with ICAM family proteins"
XX FT Binding-site
XX FT 465..474
XX FT /note= "putative cation binding site"
XX FT Binding-site
XX FT 518..527
XX FT /note= "putative cation binding site"
XX FT Binding-site
XX FT 592..600
XX FT /note= "putative cation binding site"
XX FT Region
XX FT 1109..1128
XX FT /note= "transmembrane region"
XX FT Domain
XX FT 1129..1161
XX FT /note= "cytoplasmic domain"
XX PN WO9517412-A1.
XX PD 29-JUN-1995.
XX XX
XX PF 21-DEC-1994; 94WO-US14832
XX XX

```

```

PR 05-AUG-1994; 94US-0286889.
PR 23-DEC-1993; 93US-0173497.
XX PA (ICOS-) ICOS CORP.
XX PI Gallatin WM, Van Der Vieren M;
XX XX WPI; 1995-240603/31.
DR N-PSDB; AAQ91712.
DR DR
XX Alpha sub-unit polypeptide of human beta 2 integrin - used to
XX identify potential antiinflammatory agents, for the treatment of
XX graft arteriosclerosis, inflammatory bowel disease, asthma, etc.
XX Claim 7; Page 82-87; 172pp; English.
XX CC A probe based on a partial cDNA clone (given in AAQ91727) of canine
XX alpha-TM1 was used to screen a human spleen cDNA library to identify
XX clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d.
XX CC The cDNA was manipulated to allow expression of recombinant alpha-d
XX subunit in COS and CHO cells.
XX SQ Sequence 1161 AA;
XX Query Match 58.0%; Score 3455; DB 16; Length 1161;
XX Best Local Similarity 59.9%; Pred. No. 9.le-279;
XX Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;
QY 5 VLLLTALTLCHGFLNDTENAMTFQENARGFGQSVVQLOGSRVWVCAPOEIVAAORGSLY 64
Db 5 VLLLSVLASVYGFNLDVEEPTIFQDAGGFGQSVVQFGRSLVVGAPLEVAANTQRLY 64
QY 65 QCDYSGSCEPIRLQVPVEAVNMISGLSLAATTSPPQLLACGPTVHQTCSNTYVKGLCF 124
Db 65 DCAATGNCQPIPLHIREAVNMISGLTLAATNGSRLLACGPTLHRVCGENSYSKGSCL 124
QY 125 LFGSNLRQOPKFPEALRGCPQEDSDIAFLIDGSGIIPHDPRRKKEFVSTVMEQLKXK 184
Db 125 LLGSRW-EITQVDPATPECPHQEMDIVFLIDGSGIDQDNFNQMGKFPVQAVMGQEGTD 183
QY 185 TLFSLMYSEEPRIHFTFKESQNNPNRSLKPKITQLLGRHTATGIRKVVRELNITNG 244
Db 184 TLFALMQSNLLKIHFTFTQRTSPQSLVDPIVQLKGLTPTATGILTVTVQLPHHKNNG 243
QY 245 ARKNAFKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVDGAFRSEKSRQELNTIA 304
Db 244 ARKSAKKILVITDQKYKDPLEYSQVIPAQAEKAGIRYVIGVGHAFQGTARQELNTIS 303
QY 305 SKPPRDHVPQVNNFEALKTIONQLREKIPAEIGTOTGSSSSSEHEMSQEGFSAAITNSGP 364
Db 304 SAPPQDHVPKVDNFAALGSIQKLOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDL 363
QY 365 LLSTVGSYDWDAGGVFLYTSKESKSTFINMTVDSDNDVAVLYGAAAAILLRNRVQSLVLGAP 424
Db 364 FLGAVGSFWSGATFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVQNLVLGAP 423
QY 425 RYQHIGLVAMPFRQNTGMESNANVAGTQITQYFGASLCSVDVDSNGSTDLVLIGAPHYYE 484
Db 424 RYQHTGKAVIFTOVSRQWRKKAETVGTQIGSYFGASLCSVDVDSGSTDLLIGAPHYYE 483
QY 485 QTRGGQSVCPPLPRGRARWQCDVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVAIGAP 544
Db 484 QTRGGQSVCPPLPRGRQVQWQCDVAVLRGEQHPWGRFGAALTVLGVDNEDKLIDVAIGAP 543
QY 545 GEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLOYFGQSLSGQDLTMDGLVDLTV 604
Db 544 GEQENRGAVYLFHGTSGSGISPSHSQRTAGSKLSRLOYFGQSLSGQDLTMDGLMDLAV 603
QY 605 GAQGHVLLRSQPLVRKRAIMEFNPVARNVFECDQVVKGEAGEVRCVLCVOKSTRD 664
Db 604 GARGQVLLRLSLPLVVKGVAMRFPSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLD 663
QY 665 RLREGIQSVTVYDUALDSDGRPHSRVAVFNETKNSTRQTVLGLTQTCETLKLQUPNCIE 724

```

```
Db 664 QL--GDIOSSVRFDLADPGRLTSRAIFNETKNPTLTKRKTGLGHCETLKLKLLPDCVE 721
Qy 725 DPVSPVLRNLSVGTPLSFAFGNLRPLVLAEDAQRLETFALFPFKNCGNDNICODDLISIT 784
Db 722 DWSPFIHLNLSVLRPIPSQNLPRVLAAGSODLFTASLPFKNCGQDGLCEGDLGVT 781
Qy 785 FSPMSLDCLVGGPREFNVTVRNDDGDSVRTQVTFPPFLDLRYKRVSTLQNRQSRWS 844
Db 782 LSPGLQTLTVGSSLELNVITVWVWAGDSVTVVSLYYPAGLSHRVRSGAQKPHQSAL 841
Qy 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
Db 842 RLACETV-PTDEG-LRSSRCSVNHPIFHEGSGNGTFIVTFDVSFKATILGDRMLRASASS 899
Qy 905 ENNPRTNKTFFOLEPLPKYAVYVWVTSYGSTKYNP-TASENTRVMQHOYQVNLQ 963
Db 900 ENNKASSKATFOLEPLPKYAVYVWVTSYGSTKYNP-TASENTRVMQHOYQVNLQ 959
Qy 964 RSLPISLVLVPLVRLNQTVIDRPOVTFSENLSSCTCHERLPSHSDPLAELRKAPVVC 1023
Db 960 RDLAISINFWVLLNGVAVWVWVMEAPSQSL--PCVSEKPPQHSDFLTQISRPMDC 1017
Qy 1024 STAVQRIQCDDIPFGIIEFNATLGNLSDWYIKTSHNLLIVSTAEILFNDSVFTLL 1083
Db 1018 STADCLQFRCDVPSPVSOEELDFTLKGNLSFGWVRETLLQKVLVSVVAEITFDTSVSQL 1077
Qy 1084 PQGAFVRSQETKVEPEVFNPLPLVGVSSVGLLLALTAALYKLGFFKRYQKDMMS 1143
Db 1078 PQGEAFMAQMEVLEDEVNAIPIMGSSVGLLLALATLTLKLGFFKRYKEMLE 1137
Qy 1144 E 1144
Db 1138 D 1138
```

RESULT 13

AAW23049
ID AAW23049 standard; Protein; 1161 AA.

XX AC AAW23049;

XX DT 24-FEB-1998 (first entry)

XX DE Human beta 2 integrin alpha d subunit.

XX KW Beta 2 integrin alpha d subunit; human; cell migration;
cell adhesion; phagocytosis; diabetes; atherosclerosis;
multiple sclerosis; asthma; psoriasis; lung inflammation;
acute respiratory distress syndrome; rheumatoid arthritis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 17..1108 /label= Extracellular_domain

FT Domain 1109..1128 /label= Transmembrane_domain

FT FT /note= "homologous to the human CD11c transmembrane region"

FT Domain 1129..1161 /label= Cyttoplasmic_domain

FT Domain 150..352 /note= "region homologous to the I (insertion)

FT FT domain common to CD11a, CD11b and CD11c"

XX PN WO9731099-A1.

XX XX 28-AUG-1997.

XX XX 24-FEB-1997;

XX XX 22-FEB-1996;

XX PR 96US-0605672.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin WM, Van Der Vieren M;

XX XX WPI; 1997-435154/40.

DR N-PSDB; AAT79220.

XX XX Hybridoma 199M and antibody secreted by it - specific for new rat
beta2 integrin subunit, useful to detect subunit in cells and
modulate its activity

XX XX Example 5; Page 116-120; 222pp; English.

XX CC This polypeptide comprises a novel human beta 2 integrin subunit,
designated alpha d. Its sequence was deduced from a cDNA clone
(see AAT79220) isolated from a spleen cDNA library. Alpha d is
involved in cell migration, phagocytosis and cell-cell interaction.
Recombinant alpha d polypeptides can be expressed in transformed
host cells for use in assays for identifying antibodies or other
compounds that modulate alpha d activity or which modulate the
interaction between alpha d and a ligand, for treating or
preventing diseases in which macrophages are implicated. Treatment
is applicable to disease states in which alpha d binding, or
localised accumulation of cells which express alpha d, is
implicated such as type 1 diabetes, atherosclerosis,
multiple sclerosis, asthma, psoriasis, lung inflammation, acute
respiratory distress syndrome and rheumatoid arthritis.

XX SQ Sequence 1161 AA;

Query Match 59.0%; Score 3455; DB 18; Length 1161;
Best Local Similarity 59.9%; Pred. No. 9.1e-279;
Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;

Qy 5 VLLLTALTLCGHNLDNTENAMTFQENARGGQSVVQLQSGSRVVGAPQEIIVAANQSGLY 64
Db 5 VLLLSVLASYHGHNLDVEEPTIFQEDAGGFGQSVVFGGSRVVGAPLEVVAANQTRL 64
Qy 65 QCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTPPOLLAGCTPTVHVCSENTVYKGLCF 124
Db 65 DCAATGMCQPIPLHIREAVNMSLGLTLAATNGSRLLACGPTLHRVCGENSKGSL 124
Qy 125 LFGSNLRQOPQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKPEFVSTMOLKSK 184
Db 125 LLGSRW-EIIQTVPDATPECPQEBMDIVFLIDGSGSDQDNFNQMGKFPVQVMOFGETD 183
Qy 185 TLFSLMQYSEEFRIHFTKFEQNNPNSRLVKPITOLLGRTHATGIRKVVRELPNITNG 244
Db 184 TLFALMOYSNLLKIHFTTQFTSPSQSLVDPIVQLKGLTFTATGILTVVTPHKKNG 243
Qy 245 ARKNAFKILVITDGEKFGDPLGYEDVPIPEADREGVIRYVGVGDAFSEKSRQBLNTIA 304
Db 244 ARKSANKILVITDQKQKDPLEYSDDVIPQAEKAGIIRYALGVGHAFQGPARTARQELNTIS 303
Qy 305 SKPPRDHVFQVNNFEALKTIQNLRKFIPIEGTQSGSSSEFHEHMSQEGSAAITSNGP 364
Db 304 SAPPQDHFVKVDNFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGSTALTMGDL 363
Qy 365 LLSTVGSVDWAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAAILLRNVQSILVGLAP 424
Db 364 FLGAVGFSFGSGGAFLYPPNMSPTFINNSQENVDNRDLSYLGSTELWLKVGQNLVLGAP 423
Qy 425 RYQHTGLVMPFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVE 484
Db 424 RYQHTGKAVIFTQVSRQWRKKAETVGTQIGSVFGASLCSVDVDSNGSTDVLIGAPHYVE 483
Qy 485 QTRGGQSVSVCLPRGQARWQCDAYLYGEQGGQPGWRFGAALTVLGADVNGDKLTDVAIGAP 544
Db 484 QTRGGQSVSVCLPRGQARWQCDAYLYGEQGGQPGWRFGAALTVLGADVNGDKLTDVAIGAP 543
Qy 545 GEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSRPLQYFGQSLSGGQDLTWDGLVDLTV 604

```
Db 544 GEQNRGAVYLFHGAESGIGSPHSQRIASSQLSPRLQVFGQALSGGQDLTDGGLMDLAV 603
Qy 605 GAQGHVLLRSQVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRD 664
Db 604 GARGQVLLRLSLPVLKGVAMFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLD 663
Qy 665 RLREBQIQSVTYDIALDSGRPHSRVAFNETKXSTRQTVGLGTQTCETLKLQLPNCIE 724
Db 664 QL--GDIQSSVRFDLALDGRITSRAIFNETKXNPTLTRKTLGLGIHCETLKLPLDCVE 721
Qy 725 DVPSPVLRLNLSLVTPLSAFNGLRPLVLAEDAQRLEFALRPFKNCNDNICOGLSIT 784
Db 722 DVSPPIILHLNLSLVREPIPSQNLRLPVLAVGSQDLFTASLRFKNCQDGLCEGLDGT 781
Qy 785 FSNLSDCLVWGGPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRW 844
Db 782 LSFSGQLTLTVGSSLELVIVTMNAGEDSYGVTVSLYYPAGLSHRRVSGAQKPHQSAL 841
Qy 845 RIACESASTESVGALKSTSCSINHPIPPENSEVTNFTFDVDSKASLGNKLLKANVTS 904
Db 842 RIACETV--PTEDEG--LRSRCSVNHPIFHEGSGNGTFIVTFDVSYKATLGDRLMLRASASS 899
Qy 905 ENNMERTNKTETOLELPVKYAVYVMTSHGVSTKYLN--TASENTSRVMQHOYQVSNLQ 963
Db 900 ENNKASSKATFQLELPVKYAVYTMISROESTKYFNFATSDEKXKKAHRYRNNLSQ 959
Qy 964 RSLPISLVLVPLVRLNQTWIDRPQVTFSENLSSTCHTKERLPSHSDFLAEILRKAPVWNC 1023
Db 960 RLALISINFWPVLNGVAVWDMVMEAPSQL--PCVSEKPPQSHDELQISRSPMLDC 1017
Qy 1024 STAVCORIQCDIPFGIQBEFNATLKNLSFOWYIKTSHNHLIVSTAEILPNDVSFTLL 1083
Db 1018 STADCLQFRCDVPFSFVQBELDTLKNLSFGWVRETLQKKVLVSVVAEITPDTSVYSOL 1077
Qy 1084 PCOGAFVRSQETKVEPEVPNPLPLIVGSSVGGILLILALITAAALKYKGFPRQYKDWMS 1143
Db 1078 PCOEAFMRQAMWLEEDVYNAIPIIMGSSVGAALLLALITATLYKLGFFKRYKEMLE 1137
Qy 1144 E 1144
Db 1138 D 1138

RESULT 14
AAW72825
ID AAW72825 standard; Protein; 1161 AA.
XX AC AAW72825;
XX DT 19-JAN-1999 (first entry)
XX DE Human alpha-d.
XX KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
XX KW immunoglobulin; fusion protein; binding molecule; antibody;
XX KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
XX KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
XX KW acute respiratory distress syndrome.
XX OS Homo sapiens.
XX PN US5831029-A.
XX PD 03-NOV-1998.
XX PF 07-JUN-1995; 95US-0482293.
XX PR 07-JUN-1995; 95US-0482293.
XX PR 23-DEC-1993; 93US-0173497.
XX PR 05-AUG-1994; 94US-0286889.
XX PR 21-DEC-1994; 94US-0362652.
XX PA (ICOS-) ICOS CORP.

XX PI Gallatin WM, Van Der Vieren M;
XX DR WPI; 1998-609318/51.
XX DR N-PSDB; AAV67281.
XX PT Antibodies specific for beta-2 integrin alpha-subunit d - useful in
XX PT immunohistochemical analysis
XX PS Example 5; Column 61-66; 106pp; English.
XX CC The present sequence represents human alpha d. The present invention
XX CC describes: (1) an antibody that specifically binds alpha d, which
XX CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
XX CC antibody as in (1); (3) an anti-idiotypic antibody specific for the
XX CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
XX CC monoclonal antibody of (2). Antibodies specific for alpha d can be
XX CC used in immunohistochemical analysis to localise alpha d to subcellular
XX CC compartments or individual cells within tissues. Substances that modulate
XX CC alpha d binding (which may include antibodies) can be used to treat
XX CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
XX CC inflammation, acute respiratory distress syndrome or arthritis.
XX SQ Sequence 1161 AA;
Query Match 58.0%; Score 3455; DB 19; Length 1161;
Best Local Similarity 59.9%; Pred. No. 9.1e-279;
Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;
Qy 5 VLLLTALTLCGHNLDNENAMTFQENARGFQSVVQVQSGRVVVGAPQEIIVAAHQSGSLY 64
Db 5 VLLLSVLASYHGFNLDBEETIFQEDAGGFGQSVVQVQSGRLVVGAPLEVVAAHQGTGLY 64
Qy 65 QCDYSTGSCBRLQVPEAVNMSLGLSLAATTPPOLLACGPTVHOCTSENTRYKGCFCF 124
Db 65 DCAATGMCQPIPHIRPEAVNMSLGLTAASTNGSRLACGPTLHRVCGENSGKSGCL 124
Qy 125 LFGNLRQOPKPFPEARLGCPEQSDIAFLIDGSGSIIPHPFRMKFEVSTVMEQKKSK 184
Db 125 LLSRW--EIIQTVDPATPECPHQMEDIIVFLIDGSGSIDQDNFMQKGFQVAMQFEGTD 183
Qy 185 TLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPTQLLGRTHRTATGIRKRVRELPNITNG 244
Db 184 TLFALMQYSNLLKHFTFTQRTSPSQSLVDPIVQLKGLTFTATGILTIVVTQLPHKNG 243
Qy 245 ARKNAFKILVITDGEKFGDPLGVEDVITPEADREGVIRYVIGDAPFSEKSRQELNTIA 304
Db 244 ARKSAKILVITDQKYKDPLEYSDVIPQAEKAGIRIYAGVGHAFQGPATARQELNTIS 303
Qy 305 SKPPRDHVFQVNFPEALKTIQNLREKIFAIEGTQSGSSFEHEMSQEGFSAATISNGP 364
Db 304 SAPQDHFVKVDNFAALGSIQKQLEKIYAVEGTQSRASSFQHEMSQEGFSTALTMGL 363
Qy 365 LLSTVGSVDWAGGVPLYTSKEKSTFINNTRVDSNDMDAYLGVAAILLRNRVQSLVLGAP 424
Db 364 FLGAVGSPSWSGGAFLYPPNMSPTFINNSQSNVMDRDSYLGYSTELALWKGVQNLVLGAP 423
Qy 425 RYQHIGLVAMFRONTGMHESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYVE 484
Db 424 RYQHTGKAVIFTQVSQRQMRKKAETVGTQIGSYFGASLCSVDVDSGSDTLIGAPHYVE 483
Qy 485 QTRGGQSVCPPLPRGQARMCODAVLYGECQOPMGRFCAALTVLGDVNGDKLTDVAIGAP 544
Db 484 QTRGGQSVCPPLPRGQVQWOCDAVLRGEQHPMGRFCAALTVLGDVNGDKLTDVAIGAP 543
Qy 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYPGQSLSGGQDLTDGGLMDLTV 604
Db 544 GEENRGAVYLFHGAESGIGSPHSQRIASSQLSPRLQVFGQALSGGQDLTDGGLMDLAV 603
Qy 605 GAQGHVLLRLRSQVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRD 664
Db 604 GARGQVLLRLSLPVLKGVAMFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLD 663
```


Qy	545	GEENRGAVYLFHGTSGSISPSHSQRIAGSKLSPRLQYFCQSLSGGQDLTMDGLVDLTV	604
Db	544	GEENRGAVYLFHGTSGSISPSHSQRIAGSKLSPRLQYFCQSLSGGQDLTMDGLMDLAV	603
Qy	605	GAQGHVLLLRQPVLRVKAIMEFNPREFVARNVFCNDQVVKGEAGEVVRVCLHVOKSTRD	664
Db	604	GARGOVLLLRSLPVLKVGAVMRFSVEVEKAVYRCWEKPSALEAGDATVCLTIQKSSLD	663
Qy	665	RLREGQIQSVVYDLDLSDGRPHRAVFNETKNSRRQTVLGLTQTCETLKLQLPNCIE	724
Db	664	QL--GDIOSSVRFDLALDPGLTSRAIENETKNPTLFRKTLGLGIHCETLKLPLDCVE	721
Qy	725	DPVSPVLRLNFSLVGTPLSAFGLNRPVLAEDAQRFTALFPFKKNCNDNICODDLSIT	784
Db	722	DWSPILHLNFSLVREIPSPQNLRLPVLAVGSQDLFTASLPFEKNCQDGLCEGLGVT	781
Qy	785	FSFMSLDCLVGGPREFNVTVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQORSW	844
Db	782	LSFSGLOTLTVGSSLELNVIVTMNAGEDSYGTVVSLYPAGLSHRRVSGAQKQPHOSAL	841
Qy	845	RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS	904
Db	842	PLACETV-PTDEG-LRSRCSVNHPIPHESNGTFIVTFDVSYKATLGDRLMLRASASS	899
Qy	905	ENNMPTNKTBFQLELPVKYAVYMYVTSHGYSTKYNF-TASENTSRVMQHOYQVSNLQ	963
Db	900	ENNKASSKATFQLELPVKYAVYTMISQEESTKYFNFATSDEKMKKEAHRVYVNNLSQ	959
Qy	964	RLPISLVLPVRLNQTIVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVC	1023
Db	960	RDLAISINFWPVLINGVAVWVMEAPQSLL--PCVSEKPPQHSDFLTQISRSPMLDC	1017
Qy	1024	SIACQRIQCIPPFQIOEFNATLKGNLSDWIKTSHNHLLIVSTAELLFNDVSFTLL	1083
Db	1018	SIADCLQPCDVPSFVQBEJLDFTLKGNLSFGWVRETLOKKVLVVSVAEITFDTSVYSQL	1077
Qy	1084	PGQAFVRSQTETKVEPFPVNPPLIYGSSVGGLLLALITAALYKLGFFKRYQKMMMS	1143
Db	1078	PGQAFMEAQMEVLEDEEVYNAIPIINGSSVGALLLALITATLYKLGFFKRYKEMLE	1137
Qy	1144	E	1144
Db	1138	D	1138

Search completed: November 25, 2003, 14:16:22
Job time : 40.8086 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:20:45 ; Search time 26.4941 Seconds
(without alignments)
8026.803 Million cell updates/sec

Title: US-09-902-481b-1

Perfect score: 5956

Sequence: 1 MALRVLLLTALTLCHGFNLD.....FKRQYKDMMEGGPGAEPO 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 18443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5956	100.0	1153	9	US-09-350-259-3
2	5956	100.0	1153	11	US-09-902-481A-1
3	5956	100.0	1153	11	US-09-891-943-3
4	5956	100.0	1153	15	US-10-144-259-30
5	5956	100.0	1153	15	US-10-207-655-176
6	5940.5	99.7	1152	10	US-09-945-265-4
7	5868	98.5	1137	11	US-09-902-481A-6
8	5862	98.4	1137	11	US-09-902-481A-5
9	5852	98.3	1137	11	US-09-902-481A-4
10	5839	98.0	1137	11	US-09-902-481A-3
11	3503	58.8	1163	12	US-10-116-275-204
12	3480	58.4	1163	9	US-09-350-259-4
13	3480	58.4	1163	11	US-09-891-943-4
14	3455	58.0	1161	9	US-09-350-259-2
15	3455	58.0	1161	11	US-09-891-943-2

16 3439.5 57.7 9 US-09-350-259-99 Sequence 99, Appl
17 3439.5 57.7 11 US-09-891-943-99 Sequence 99, Appl
18 3270.5 54.9 9 US-09-350-259-55 Sequence 55, Appl
19 3270.5 54.9 11 US-09-891-943-55 Sequence 55, Appl
20 3264 54.8 11 US-09-350-259-53 Sequence 53, Appl
21 3264 54.8 11 US-09-891-943-53 Sequence 53, Appl
22 3251.5 54.6 11 US-09-350-259-37 Sequence 37, Appl
23 3251.5 54.6 11 US-09-891-943-37 Sequence 37, Appl
24 3243 54.4 11 US-09-350-259-46 Sequence 46, Appl
25 3243 54.4 11 US-09-891-943-46 Sequence 46, Appl
26 1563 26.2 11 US-09-945-265-2 Sequence 2, Appl
27 1229.5 20.6 9 US-09-350-259-103 Sequence 103, Appl
28 1229.5 20.6 494 11 US-09-891-943-103 Sequence 103, Appl
29 1189.5 20.0 413 9 US-09-350-259-101 Sequence 101, Appl
30 1189.5 20.0 413 11 US-09-891-943-101 Sequence 101, Appl
31 1163 19.5 11 US-10-177-550-2 Sequence 2, Appl
32 1108.5 18.6 11 US-09-984-130-103 Sequence 103, Appl
33 1108.5 18.6 11 US-09-836-353A-103 Sequence 103, Appl
34 1102.5 18.5 11 US-09-984-130-35 Sequence 35, Appl
35 1102.5 18.5 11 US-09-836-353A-35 Sequence 35, Appl
36 1087.5 18.3 11 US-10-160-354-2 Sequence 2, Appl
37 1042 17.5 216 12 US-09-795-872-5 Sequence 5, Appl
38 1036 17.4 240 15 US-10-102-806-453 Sequence 453, Appl
39 987 16.6 1034 11 US-09-984-130-43 Sequence 43, Appl
40 987 16.6 1034 12 US-09-836-353A-43 Sequence 43, Appl
41 979 16.4 191 11 US-09-805-354-1 Sequence 1, Appl
42 979 16.4 191 15 US-10-144-259-1 Sequence 1, Appl
43 875 14.7 707 9 US-09-764-870-313 Sequence 313, Appl
44 875 14.7 707 15 US-10-125-540-313 Sequence 313, Appl
45 850 14.3 223 12 US-10-002-631C-197 Sequence 197, Appl

ALIGNMENTS

RESULT 1

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 100.0%; Score 5956; DB 9; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLGSRVVVVCAPQEIIVANOR 60
DB 1 MALRVLLLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLGSRVVVVCAPQEIIVANOR 60
QY 61 GSLYQCDYSGTSGCEPIRLQVPVEAVNMISLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120

Db 61 GSYLQCDYSGCEPIRLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVK 120
Qy |||||
Db 121 GLCLFGLNLRQOPQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
Qy |||||
Db 121 GLCLFGLNLRQOPQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
Qy |||||
Db 181 KKSXTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFN 240
Qy |||||
Db 181 KKSXTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFN 240
Qy |||||
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKROEL 300
Qy |||||
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKROEL 300
Qy |||||
Db 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAEGTQTGSSSSFEHMSQEGFSAIT 360
Qy |||||
Db 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAEGTQTGSSSSFEHMSQEGFSAIT 360
Qy |||||
Db 361 SNGPLLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Qy |||||
Db 361 SNGPLLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Qy |||||
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Qy |||||
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Qy |||||
Db 481 HYYEOTRGQGVSVCPPLRGQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Qy |||||
Db 481 HYYEOTRGQGVSVCPPLRGQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Qy |||||
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600
Qy |||||
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600
Qy |||||
Db 601 DLTUGAQHVLRLRQOPVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQK 660
Qy |||||
Db 601 DLTUGAQHVLRLRQOPVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQK 660
Qy |||||
Db 661 STRDLREGIOISVTVYDLALDSGRPHSRVNETKNSTRQTVLGTOTCETLKLQLP 720
Qy |||||
Db 661 STRDLREGIOISVTVYDLALDSGRPHSRVNETKNSTRQTVLGTOTCETLKLQLP 720
Qy |||||
Db 721 NCIEDPVPVILRLNFSLVGTPLSAFAGNLRPVLAEDAORLFTALPFPKNCNDNICODD 780
Qy |||||
Db 721 NCIEDPVPVILRLNFSLVGTPLSAFAGNLRPVLAEDAORLFTALPFPKNCNDNICODD 780
Qy |||||
Db 781 LSITFSFMSLCLVVGGRPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR 840
Qy |||||
Db 781 LSITFSFMSLCLVVGGRPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR 840
Qy |||||
Db 841 QRSWRLACESASSTEVSGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy |||||
Db 841 QRSWRLACESASSTEVSGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy |||||
Db 901 NVTSENNMPTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQYQVSN 960
Qy |||||
Db 901 NVTSENNMPTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQYQVSN 960
Qy |||||
Db 961 LQGRSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAEURKAPV 1020
Qy |||||
Db 961 LQGRSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAEURKAPV 1020
Qy |||||
Db 1021 VNCSTAVCORIOCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIIIVSTAEILFNSVF 1080
Qy |||||
Db 1021 VNCSTAVCORIOCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIIIVSTAEILFNSVF 1080
Qy |||||
Db 1081 TLLPGQAFVRSQETETKVEPFPVNPPLPLIVGSSVGGLLLLALITAAALYKGLFFRQYKD 1140
Qy |||||
Db 1081 TLLPGQAFVRSQETETKVEPFPVNPPLPLIVGSSVGGLLLLALITAAALYKGLFFRQYKD 1140
Qy |||||
Db 1141 MMSEGGPPGAEPQ 1153
Qy |||||

Db 1141 MMSEGGPPGAEPQ 1153
RESULT 2
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/REF/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 100.0%; Score 5956; DB 11; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTLCGHNLTENAMTFQENARGFGQSVQLOQSRVVVVGAPQEIIVANQR 60
Db 1 MALRVLLLTALTLCGHNLTENAMTFQENARGFGQSVQLOQSRVVVVGAPQEIIVANQR 60
Qy 61 GSYLQCDYSGCEPIRLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVK 120
Db 61 GSYLQCDYSGCEPIRLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVK 120
Qy 121 GLCLFGLNLRQOPQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
Db 121 GLCLFGLNLRQOPQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL 180
Qy 181 KKSXTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFN 240
Db 181 KKSXTLFSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKROEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKROEL 300
Qy 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAEGTQTGSSSSFEHMSQEGFSAIT 360
Db 301 NTIASKPRDHVQVNNFEALKTIONLREKIFAEGTQTGSSSSFEHMSQEGFSAIT 360
Qy 361 SNGPLLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Db 361 SNGPLLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Qy 481 HYYEOTRGQGVSVCPPLRGQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYYEOTRGQGVSVCPPLRGQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQSGLSGGQDLTMDGLV 600

```
QY 601 DLTGAGQGHVLLRSQPLVRKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTGAGQGHVLLRSQPLVRKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLREGQIQSVVYTDIALDSGRPHSRAVFNKTSRRTQVGLTQTCTETLKLQLP 720
DB 661 STRDLREGQIQSVVYTDIALDSGRPHSRAVFNKTSRRTQVGLTQTCTETLKLQLP 720
QY 721 NCIEDPVSPIVLRNFSLVGTPLSFGNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
DB 721 NCIEDPVSPIVLRNFSLVGTPLSFGNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
QY 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNR 840
DB 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNR 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNMPRTNKTETFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOQVSN 960
DB 901 NVTSENNMPRTNKTETFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOQVSN 960
QY 961 LGORSPLISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGORSPLISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCISAVCORIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSNNHLLIVSTABILFNDVSF 1080
DB 1021 VNCISAVCORIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSNNHLLIVSTABILFNDVSF 1080
QY 1081 TLLPGGAFVRSQETKVPPEFVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKD 1140
DB 1081 TLLPGGAFVRSQETKVPPEFVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKD 1140
QY 1141 MMSEGGPPGAEPQ 1153
DB 1141 MMSEGGPPGAEPQ 1153
```

RESULT 3

```
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3
```

```
Query Match 100.0%; Score 5956; DB 11; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEI VAAHQ 60
DB 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEI VAAHQ 60
QY 61 GSVYQCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATSPPLLACGPTVHTCSENTYVK 120
DB 61 GSVYQCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATSPPLLACGPTVHTCSENTYVK 120
QY 121 GLCFLFGSNLRQOQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKSFVSTVMEQL 180
DB 121 GLCFLFGSNLRQOQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKSFVSTVMEQL 180
QY 181 KKSXTLFLSMQYSEEFRIHFTFKBFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KKSXTLFLSMQYSEEFRIHFTFKBFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIYRIVGVGDAPFRSEKROEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIYRIVGVGDAPFRSEKROEL 300
QY 301 NTIASKPRDRHVQVNNFEALKTIQNLREKIFAEGTQTCSSSFHEMSEOGESAAIT 360
DB 301 NTIASKPRDRHVQVNNFEALKTIQNLREKIFAEGTQTCSSSFHEMSEOGESAAIT 360
QY 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
DB 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASICSDVDNSGSTDVLVIGAP 480
DB 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASICSDVDNSGSTDVLVIGAP 480
QY 481 HYYETRGQGVSVCPPLPRGQARWQCDVLYGEOQPGWRFGAALTVLGDVNGDKLTQVA 540
DB 481 HYYETRGQGVSVCPPLPRGQARWQCDVLYGEOQPGWRFGAALTVLGDVNGDKLTQVA 540
QY 541 IGAPCEENRGAVALYFHGTSGSGISPSHSQIRTAGSKLSPRLQYFCQSLSGGQDLTMDGLV 600
DB 541 IGAPCEENRGAVALYFHGTSGSGISPSHSQIRTAGSKLSPRLQYFCQSLSGGQDLTMDGLV 600
QY 601 DLTGAGQGHVLLRSQPLVRKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTGAGQGHVLLRSQPLVRKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLREGQIQSVVYTDIALDSGRPHSRAVFNKTSRRTQVGLTQTCTETLKLQLP 720
DB 661 STRDLREGQIQSVVYTDIALDSGRPHSRAVFNKTSRRTQVGLTQTCTETLKLQLP 720
QY 721 NCIEDPVSPIVLRNFSLVGTPLSFGNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
DB 721 NCIEDPVSPIVLRNFSLVGTPLSFGNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
QY 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNR 840
DB 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNR 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNMPRTNKTETFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOQVSN 960
DB 901 NVTSENNMPRTNKTETFOLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOQVSN 960
QY 961 LGORSPLISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGORSPLISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCISAVCORIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSNNHLLIVSTABILFNDVSF 1080
DB 1021 VNCISAVCORIQCDIPFFGQIEEFNATLKGNSLFDWYIKTSNNHLLIVSTABILFNDVSF 1080
QY 1081 TLLPGGAFVRSQETKVPPEFVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKD 1140
```

Db 1081 TLLPGQAFVRSQTETKVEPEVNPPLIVGSSVGGLLALLALITAAALYKLGFFKQYKD 1140
Qy 1141 MMSEGGPPGAPQ 1153
Db 1141 MMSEGGPPGAPQ 1153
RESULT 4
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30
Query Match 100.0%; Score 5956; DB 15; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRVLLLTALTALCHGNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIWAANQR 60
Db 1 MALRVLLLTALTALCHGNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIWAANQR 60
Qy 61 GSLYQCDYSTGSCSPIRLQVPVEAVNMVSLGLSLAATSPPLLACGPTVHQTCSNTYVK 120
Db 61 GSLYQCDYSTGSCSPIRLQVPVEAVNMVSLGLSLAATSPPLLACGPTVHQTCSNTYVK 120
Qy 121 GLCLFGLNLRQOPKFEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQL 180
Db 121 GLCLFGLNLRQOPKFEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQL 180
Qy 181 KKSFTLSLMQYSEFRHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KKSFTLSLMQYSEFRHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFVLVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFVLVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQEL 300
Qy 301 NTIASKPRDRHVQVNNFEALKTIONLREKIFAIEGTQTGSSSFEHMSQEGESAIT 360
Db 301 NTIASKPRDRHVQVNNFEALKTIONLREKIFAIEGTQTGSSSFEHMSQEGESAIT 360
Qy 361 SNGPLLLSTVGSVDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV 420
Db 361 SNGPLLLSTVGSVDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Qy 481 HYYEOTRGQVSVCPPLRGQRARWOCDAVLVGEQOPWGRGAALTIVLGDVNGDKLTDVA 540
Db 481 HYYEOTRGQVSVCPPLRGQRARWOCDAVLVGEQOPWGRGAALTIVLGDVNGDKLTDVA 540
Qy 541 IGAPCEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600

Db 541 IGAPCEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Qy 601 DLTVGAGQHVLLRSQPVLRVKAIMFNPREVARNVFECDQVVKKEAGEVRCVCLHVQK 660
Db 601 DLTVGAGQHVLLRSQPVLRVKAIMFNPREVARNVFECDQVVKKEAGEVRCVCLHVQK 660
Qy 661 STRDLREGQIQSVWYTDALDSGRPHSRVAFNETKNSTRROTQVGLGLTQTCETLKLQLP 720
Db 661 STRDLREGQIQSVWYTDALDSGRPHSRVAFNETKNSTRROTQVGLGLTQTCETLKLQLP 720
Qy 721 NCIEDPVSPVLRNLSVGTPLSAFGNLRPVLAEADAQRLFTALFPFKNCGNDNICQDD 780
Db 721 NCIEDPVSPVLRNLSVGTPLSAFGNLRPVLAEADAQRLFTALFPFKNCGNDNICQDD 780
Qy 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNR 840
Db 781 LSITFSFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNR 840
Qy 841 QRSWRLACESASSTEVSGLKSTSCSINHPIFFENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWRLACESASSTEVSGLKSTSCSINHPIFFENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNMPTNKTEFQELPVLKYAVMVVTVSHGVSTKYLNFTASENTSRVMOHQYVSN 960
Db 901 NVTSENNMPTNKTEFQELPVLKYAVMVVTVSHGVSTKYLNFTASENTSRVMOHQYVSN 960
Qy 961 LGORSUPISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Db 961 LGORSUPISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Qy 1021 VNCISIAVCQRIQCDIPFGIOEEFNATLKGNLSPDWYIKTSHNHLITVSTAEILFNDSVF 1080
Db 1021 VNCISIAVCQRIQCDIPFGIOEEFNATLKGNLSPDWYIKTSHNHLITVSTAEILFNDSVF 1080
Qy 1081 TLLPGQAFVRSQTETKVEPEVNPPLIVGSSVGGLLALLALITAAALYKLGFFKQYKD 1140
Db 1081 TLLPGQAFVRSQTETKVEPEVNPPLIVGSSVGGLLALLALITAAALYKLGFFKQYKD 1140
Qy 1141 MMSEGGPPGAPQ 1153
Db 1141 MMSEGGPPGAPQ 1153
RESULT 5
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176
Query Match 100.0%; Score 5956; DB 15; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRVLLLTALTALCHGNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIWAANQR 60
Db 1 MALRVLLLTALTALCHGNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIWAANQR 60
Qy 61 GSLYQCDYSTGSCSPIRLQVPVEAVNMVSLGLSLAATSPPLLACGPTVHQTCSNTYVK 120

```
Db 61 GSYLQCDYSTGSCPEIRLQVPVEAVNMVSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
Qy 121 GLCFLFGSNLRQOPKQFPALRCCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOL 180
Db 121 GLCFLFGSNLRQOPKQFPALRCCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOL 180
Qy 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPFSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPFSEKSRQEL 300
Qy 301 NTIASKPPRDHVQVNNFEALKTIONQLREKIFAIEGTOTGSSSSFEHEMSQEGFSAAT 360
Db 301 NTIASKPPRDHVQVNNFEALKTIONQLREKIFAIEGTOTGSSSSFEHEMSQEGFSAAT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYVEOTRGGQSVSCPILPRGORARWQCDVLYGEGOPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYVEOTRGGQSVSCPILPRGORARWQCDVLYGEGOPWGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKKEAGEVRVCLHVQK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKKEAGEVRVCLHVQK 660
Qy 661 STRDRREGOIQSVVTVYDALDSGRPHSRAVFNKTNSTRROTQVGLTQTCETLKLQIP 720
Db 661 STRDRREGOIQSVVTVYDALDSGRPHSRAVFNKTNSTRROTQVGLTQTCETLKLQIP 720
Qy 721 NCIEDPVSPIVLRNLSVGTPLSAFGNLRPVLAEADAQRLFTALFPFKNCGNDNICQDD 780
Db 721 NCIEDPVSPIVLRNLSVGTPLSAFGNLRPVLAEADAQRLFTALFPFKNCGNDNICQDD 780
Qy 781 LSITFSFMSLDCLVVGGRPEFNTVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRS 840
Db 781 LSITFSFMSLDCLVVGGRPEFNTVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRS 840
Qy 841 ORSMRLACSSASTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 ORSMRLACSSASTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENMPRTNKTQELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 960
Db 901 NVTSENMPRTNKTQELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 960
Qy 961 LGQRSPLISLVLPVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKPV 1020
Db 961 LGQRSPLISLVLPVRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKPV 1020
Qy 1021 VNCISIAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVF 1080
Db 1021 VNCISIAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVF 1080
Qy 1081 TLLPGOGAFVRSOTETKVEPFPVNPPLIVTGVSSVGLLILLALITAAALKYLGFFKQYKD 1140
Db 1081 TLLPGOGAFVRSOTETKVEPFPVNPPLIVTGVSSVGLLILLALITAAALKYLGFFKQYKD 1140
Qy 1141 MMSEGGPPGAEPQ 1153
Db 1141 MMSEGGPPGAEPQ 1153
```

RESULT 6

```
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4
```

```
Query Match 99.7%; Score 5940.5; DB 10; Length 1152;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
Qy 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARFGQSVVQLQGSRVVVGAPQEI1VAANQR 60
Db 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARFGQSVVQLQGSRVVVGAPQEI1VAANQR 60
Qy 61 GSYLQCDYSTGSCPEIRLQVPVEAVNMVSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
Db 61 GSYLQCDYSTGSCPEIRLQVPVEAVNMVSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
Qy 121 GLCFLFGSNLRQOPKQFPALRCCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOL 180
Db 121 GLCFLFGSNLRQOPKQFPALRCCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOL 180
Qy 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KKSCTLFLSMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPFSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPFSEKSRQEL 300
Qy 301 NTIASKPPRDHVQVNNFEALKTIONQLREKIFAIEGTOTGSSSSFEHEMSQEGFSAAT 360
Db 301 NTIASKPPRDHVQVNNFEALKTIONQLREKIFAIEGTOTGSSSSFEHEMSQEGFSAAT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYVEOTRGGQSVSCPILPRGORARWQCDVLYGEGOPWGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYVEOTRGGQSVSCPILPRGORARWQCDVLYGEGOPWGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKKEAGEVRVCLHVQK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKKEAGEVRVCLHVQK 660
```

661 STRDLREGQIQSVVYDLDLSDGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLP 720
Db STRDLREGQIQSVVYDLDLSDGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLP 719
Qy NCIEDPVPVILRLNLSLVGTPLSAFGNLRVLAEDAQRFTALFPPEKNCNDNIQDOD 780
Db NCIEDPVPVILRLNLSLVGTPLSAFGNLRVLAEDAQRFTALFPPEKNCNDNIQDOD 779
Qy LSITFSFMSLCLVVGGRPRENVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840
Db LSITFSFMSLCLVVGGRPRENVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 839
Qy QRSWLACASSTEVSGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKA 900
Db QRSWLACASSTEVSGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKA 899
Qy NVTSENNMPTNKTEFQLELPLVKYAVVMVTSBGVSTKYLNTASNTSRVMQHQYQVSN 960
Db NVTSENNMPTNKTEFQLELPLVKYAVVMVTSBGVSTKYLNTASNTSRVMQHQYQVSN 959
Qy LGORSPLSLVPLVRLNQTIVMDRPOVTFSENLSSCTKTERLPKSHSDFLAELRKAPV 1020
Db LGORSPLSLVPLVRLNQTIVMDRPOVTFSENLSSCTKTERLPKSHSDFLAELRKAPV 1019
Qy VNCSTAVCQRIQCQDIPFFGIGIOEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEIIFNDSVF 1080
Db VNCSTAVCQRIQCQDIPFFGIGIOEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEIIFNDSVF 1079
Qy TLLPQOGAFVRSQTEKVEPEVPNPLPLIVGSSVGGLLLALITAALYKLGFFKQYKD 1140
Db TLLPQOGAFVRSQTEKVEPEVPNPLPLIVGSSVGGLLLALITAALYKLGFFKQYKD 1139
Qy 1141 MMSEGGPPGAEPO 1153
Db 1140 MMSEGGPPGAEPO 1152

RESULT 7

US-09-902-481A-6

; Sequence 6, Application US/09902481A

; Publication No. US2003005440A1

; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy

; APPLICANT: Shimaoka, Motomu

; APPLICANT: Shifman, Julia

; APPLICANT: Mayo, Stephen

; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

; FILE REFERENCE: A-70586-1/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/902,481A

; PRIOR FILING DATE: 2001-07-09

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 1137

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: synthetic

US-09-902-481A-6

Query Match

Best Local Similarity 98.5%; Score 5868; DB 11; Length 1137;

Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 17 FNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQEIIVAAQNRGLYQCDYSTGSCBPI 76
Db 1 FNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQEIIVAAQNRGLYQCDYSTGSCBPI 60
Qy 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCFLFGNLRQPOK 136
Db 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCFLFGNLRQPOK 120

RESULT 8

US-09-902-481A-5

; Sequence 5, Application US/09902481A

Qy 137 FPALRGCCQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTVMEQLKSKKTLFSLMQYSEEF 196
Db 121 FPALRGCCQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTVMEQLKSKKTLFSLMQYSEEF 180
Qy 197 RIHFTPKFONNPNRSLVKPITQLGRTHRTATGIRKVRRELFTNITNGARKNAFKLLVVI 256
Db 181 RIHFTPKFONNPNRSLVKPITQLGRTHRTATGIRKVRRELFTNITNGARKNAFKLLVVI 240
Qy 257 TDGEKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNLTIASKPPRDHVFQVN 316
Db 241 TDGEKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNLTIASKPPRDHVFQVN 300
Qy 317 NFALKTIQNLREKIFAIEGTQTGSSSSEFHEMSQEGFSAAITSNGLPSTVGSYDWAG 376
Db 301 NFALKTIQNLREKIFAIEGTQTGSSSSEFHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
Qy 377 GVFLYTSKXSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMPR 436
Db 361 GVFLYTSKXSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMPR 420
Qy 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL 496
Db 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL 480
Qy 497 PRGORARWQCDAYLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 556
Db 481 PRGORARWQCDAYLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 540
Qy 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGHVLLRSQ 616
Db 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGHVLLRSQ 600
Qy 617 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRLEGOIQSVVT 676
Db 601 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRLEGOIQSVVT 660
Qy 677 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQPNCEIDPVSPIVLRNLF 736
Db 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQPNCEIDPVSPIVLRNLF 720
Qy 737 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICQDLSITFSFMSLDCLVWG 796
Db 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICQDLSITFSFMSLDCLVWG 780
Qy 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Db 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Qy 857 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Db 841 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Qy 917 QLELPVKYAVVMVTSBGVSTKYLNTASNTSRVMQHQYQVSNLQGSRLPSLVFLVPV 976
Db 901 QLELPVKYAVVMVTSBGVSTKYLNTASNTSRVMQHQYQVSNLQGSRLPSLVFLVPV 960
Qy 977 RLNOTVIMDRPQVTFSENLSSCTKTERLPKSHSDFLAELRKAPVVCNCISVACQRIQCDIP 1036
Db 961 RLNOTVIMDRPQVTFSENLSSCTKTERLPKSHSDFLAELRKAPVVCNCISVACQRIQCDIP 1020
Qy 1037 PFGIQEENFATLKGNSLFDWYIKTSHNHLIIIVSTAEIIFNDSVFTLLPGQAFVRSQTEF 1096
Db 1021 PFGIQEENFATLKGNSLFDWYIKTSHNHLIIIVSTAEIIFNDSVFTLLPGQAFVRSQTEF 1080
Qy 1097 KVBEFPEVPNPLIVGSSVGGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1153
Db 1081 KVBEFPEVPNPLIVGSSVGGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1137

```

; Publication No. US 2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match 98.4%; Score 5862; DB 11; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 76
DB 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60

QY 77 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHOTSSENTYVVKGLCFLFGSNLRQOQPK 136
DB 61 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHOTSSENTYVVKGLCFLFGSNLRQOQPK 120

QY 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKPFVSTVMQOLKSKTFLFSLMQYSSEF 196
DB 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKPFVSTVMQOLKSKTFLFSLMQYSSEF 180

QY 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256
DB 181 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 240

QY 257 TDGEXFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316
DB 241 TDGEXFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 300

QY 317 NFEALKTIONQUREKIFAJEGTQGTSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
DB 301 NFEALKTIONQUREKIFAJEGTQGTSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360

QY 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAILRNRVQSLVGLGAPRYQHIGLVAMFR 436
DB 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAILRNRVQSLVGLGAPRYQHIGLVAMFR 420

QY 437 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCPPL 496
DB 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCPPL 480

QY 497 PRGQARWQCDVAVLYGEQOPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAIVLF 556
DB 481 PRGQARWQCDVAVLYGEQOPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAIVLF 540

QY 557 HGTSGSISPSHSQRIAGSKLSPRLOYFQOSLSGGQDLTMDGLVDLTVMGAQHVLRLRSQ 616
DB 541 HGTSGSISPSHSQRIAGSKLSPRLOYFQOSLSGGQDLTMDGLVDLTVMGAQHVLRLRSQ 600

QY 617 PVLRVKAIINEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGOQSVMVT 676
DB 601 PVLRVKAIINEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGOQSVMVT 660

QY 677 YDLALDSGRPHSRVAFNETKNSTRRTQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736
DB 661 YDLALDSGRPHSRVAFNETKNSTRRTQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720

; Query Match 98.3%; Score 5852; DB 11; Length 1137;
; Best Local Similarity 99.2%; Pred. No. 0;
; Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 76
DB 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60

QY 77 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHOTSSENTYVVKGLCFLFGSNLRQOQPK 136
DB 61 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHOTSSENTYVVKGLCFLFGSNLRQOQPK 120

QY 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKPFVSTVMQOLKSKTFLFSLMQYSSEF 196
DB 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKPFVSTVMQOLKSKTFLFSLMQYSSEF 180

QY 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256
DB 181 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 240

; Publication No. US 2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

; Sequence 4, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

1037 FFGIOEERFATLKGNLSPDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGOGAFVRSOTET 1096
1021 FFGIOEERFATLKGNLSPDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGOGAFVRSOTET 1080
1097 KVEPPEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFEKQYKDMMSGGPPGAPQ 1153
1081 KVEPPEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFEKQYKDMMSGGPPGAPQ 1137
```


181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRVRELNITNGARKNAFKILFL 240
Qy 257 TDEKFKGDPGLGYEDVIEADREGVIRVVGDAFRSEKSRQELNITIASPPRDHVFQVN 316
Db 241 TDEKFKGDPGLGYEDVIEADREGVIRVVGDAFRSEKSRQELNITIASPPRDHVFQVN 300
Qy 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDAG 376
Db 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDAG 360
Qy 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRYOHIGLVAMFR 436
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRYOHIGLVAMFR 420
Qy 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGGVSCVPL 496
Db 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGGVSCVPL 480
Qy 497 PRGORARWQCDVLYGSGQSPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYVLF 556
Db 481 PRGORARWQCDVLYGSGQSPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Qy 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Db 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Qy 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Db 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Qy 677 YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQPNCTIEDPVSPIVLRNLF 736
Db 661 YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQPNCTIEDPVSPIVLRNLF 720
Qy 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 796
Db 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 780
Qy 797 GPREFNVTVVRNDEGDSYRQVTFPPPLDLSYKYSTLQNRORSWRLACESASSTEV 856
Db 781 GPREFNVTVVRNDEGDSYRQVTFPPPLDLSYKYSTLQNRORSWRLACESASSTEV 840
Qy 857 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
Db 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Qy 917 QLELPVKYAVVMTSHGVSTKYNFTASENTSRVMQHYQVSNLQORSPLSILVFLVPV 976
Db 901 QLELPVKYAVVMTSHGVSTKYNFTASENTSRVMQHYQVSNLQORSPLSILVFLVPV 960
Qy 977 RLNOTVIWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
Db 961 RLNOTVIWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Qy 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGQAFVRSQTET 1096
Db 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGQAFVRSQTET 1080
Qy 1097 KVEPEVNPPLIIVGSSVGGLLLLALITAALYKLGFFKROYKMMSEGGPPGAPQ 1153
Db 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAALYKLGFFKROYKMMSEGGPPGAPQ 1137

RESULT 10

US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimadzu, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

FILE REFERENCE: A-70586-1/RFT/RMS/RMK
CURRENT APPLICATION NUMBER: US/09/902,481A
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/216,600
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 1137
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 98.0%; Score 5839; DB 11; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 17 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
Db 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
Qy 77 RLOVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFGLSGNLROPOK 136
Db 61 RLOVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFGLSGNLROPOK 120
Qy 137 FPBALRGCEQEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQLKSKKTLFSLMQYSEEF 196
Db 121 FPBALRGCEQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKKTLFSLMQYSEEF 180
Qy 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRVRELNITNGARKNAFKILFL 256
Db 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRVRELNITNGARKNAFKILFL 240
Qy 257 TDEKFKGDPGLGYEDVIEADREGVIRVVGDAFRSEKSRQELNITIASPPRDHVFQVN 316
Db 241 TDEKFKGDPGLGYEDVIEADREGVIRVVGDAFRSEKSRQELNITIASPPRDHVFQVN 300
Qy 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDAG 376
Db 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDAG 360
Qy 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRYOHIGLVAMFR 436
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRYOHIGLVAMFR 420
Qy 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGGVSCVPL 496
Db 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGGVSCVPL 480
Qy 497 PRGORARWQCDVLYGSGQSPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYVLF 556
Db 481 PRGORARWQCDVLYGSGQSPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Qy 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Db 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Qy 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Db 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Qy 677 YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQPNCTIEDPVSPIVLRNLF 736
Db 661 YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQPNCTIEDPVSPIVLRNLF 720
Qy 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 796
Db 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 780
Qy 797 GPREFNVTVVRNDEGDSYRQVTFPPPLDLSYKYSTLQNRORSWRLACESASSTEV 856


```

; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.4%; Score 3480; DB 9; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQSGSLY 64
DB 8 LLLFTALATSLGFLNLTDELTAFRVDVSAGFGDSVVQYANSVVVVGAPQKIIAANQIGGLY 67

QY 65 QCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGLCF 124
DB 68 QCGYSTGACEPIGLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRMVYLGCLF 127

QY 125 LFGSNLRQOPKQFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTVMEQLKSK 184
DB 128 LLGPT--QLTQRLPVSRCQEQDQIVFLIDGSGSISSRNPFATMNFVRAVISQFORPS 185

QY 185 TLFSLMQVSEFRTHFTKEFQNNPNRSLVKPITOLLGRTHATGTRKVVRELFTNG 244
DB 186 TQFSLMQPSNKFQTHFTFEFRRTSNPLSLASVHQLQGFYTTATATQNVVHRLPHASYG 245

; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.4%; Score 3480; DB 9; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 785 FSFPGLKSLVGSNLELNAEVMVMNDGEDSYGTTITFSHPAGLSYRYVABGQKQGLRSL 844
QY 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTS 904
DB 845 HLTG--CSAPVGSQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVGLDRLLLIANVSS 902
QY 905 ENNMPTNKTEFOLELPVKYAVYVMVTSYHGVSTKYLNFTAS--ENTSRVMOHQYOVSNLQ 963
DB 903 ENNIPRTSKTIFOLELPVKYAVYVSSHEQPTKYLNFSESEKESHVAMHRYOVNVLQ 962
QY 964 RSLPISLVFLVPLVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNC 1023
DB 963 RDLFVSIWFVWPVELNQEAVMMDVESHQPNPQLSRCSSEKIAPIAPSDFLAHIQKNPVLDC 1022
QY 1024 SIACVQRIQCDIPFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLL 1083
DB 1023 SIACCLFRCDVFSFVQEEELDTLKNLSFGVVRQILQKKVSVSVAEIIFDTSVYSQL 1082
QY 1084 PGQAFVRSQETKVEFEVFNPLPLIVGSSVGLLLALITAAALYKLGFFKQYKQDMMS 1143
DB 1083 PGQAFVRAQTITVLEKYKVHNPILVIGSSIGGLLLALITAVLYKVGFQKQYKEMME 1142
QY 1144 E 1144
DB 1143 E 1143

RESULT 13
US-09-891-943-4
; Sequence 4, Application US/09891943
; Publication No. US2003007728A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007728A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-4

Query Match 58.4%; Score 3480; DB 11; Length 1163;
Best Local Similarity 61.1%; Pred. No. 1.5e-310;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQSGSLY 64
DB 8 LLLFTALATSLGFLNLTDELTAFRVDVSAGFGDSVVQYANSVVVVGAPQKIIAANQIGGLY 67

QY 65 QCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGLCF 124
DB 68 QCGYSTGACEPIGLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRMVYLGCLF 127

QY 125 LFGSNLRQOPKQFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTVMEQLKSK 184
DB 128 LLGPT--QLTQRLPVSRCQEQDQIVFLIDGSGSISSRNPFATMNFVRAVISQFORPS 185

QY 185 TLFSLMQVSEFRTHFTKEFQNNPNRSLVKPITOLLGRTHATGTRKVVRELFTNG 244
DB 186 TQFSLMQPSNKFQTHFTFEFRRTSNPLSLASVHQLQGFYTTATATQNVVHRLPHASYG 245
```

Qy	245	ARKNAFKILVITDGEKGPDLGYEDVIPEADREGVIRYIVGVGDAPFRSEKSROELNTIA	304
Dd		: :	
Dd	246	ARRDAIKILVITDGKEGSDLYKOVIPMADAAGIIRYAIGVLAFQNRSNKELNDIA	305
Qy	305	SKPRDHVFQVNFEALKTIQNOLREKIFAIEGTQTCTSSSSSFEHMSOEQFSAATNSGP	364
Dd	306	SKPSQEHIKFVEDFDALUQIQNOLKEKIFAIEGTETISSSSFLEMAQEGBFAVTPDGP	365
Qy	365	LLSTVGSYDWAGGFVLTYSKEKSTFINWTRVDSMDNDAYLGYAAAITLRNRVQSGLVLP	424
Dd	366	VLGAVGSFTWSGGAFLYPPNMSPTFINNSENVDMRDSYLGYTELAKWKGVOSLVLP	425
Qy	425	RYOHIGLVAFRQNTGMNESNANYKGTOIGAYFGASLCSDVDBSNGSTDVLVLGAPHYYE	484
Dd	426	RYOHIGKAVFIQVSRQWRMKAEVIGTOIGSYFGASLCSDVDTDGSTDLVLGAPHYYE	485
Qy	485	QTRGGQVSVCPLPRGQBARHQCDVALYGEQGQWPGRGAALTVLGDVNGDKLTDVAIGAP	544
Dd	486	QTRGGQVSVCPLPRGNWR-RWNCDAVLYGEOGHPMGFRGAALTVLGDVNGDKLTDVVIGAP	544
Qy	545	GEEDNRCGAVLFGHTCSGSGISPHSQRIACSLKSPRLQYFGOSISGGODTLTMDGLVDLTV	604
Dd	545	GEENRGAVLFGHVLPGPSISPHSQRIAQSUSRLQYFGQALSQQGQDTDQGLVDLAV	604
Qy	605	GAGHVLILLRSQPVLRYKAIMEFNPREVANRVPECNDVOVKGEAGEVRVCLHVQKSTRD	664
Dd	605	GARGOVLLLTRPVLWVGVSQMOPPAIPRSAFECREQVVSQETLVQSNICLIYDKRSKN	664
Qy	665	RLREGIOQSVVTYDALDPGRPHRSRAVENETKNSRRQTOVLGTOTCETKLQLPNCIE	724
Dd	665	LLGSRDLQSSVTLDLALAPGRLSPRAIFQETKNSLSRVVLGKAHCENFNLLPSCVE	724
Qy	725	DPVSPVILRNFSLVGTPLSAFCGNLRPVLAEDAQRLETALPFPEKCGNDNICQDDLSIT	784
Dd	725	DSVPIITLRLNFTLVGRKPFLAFLARNLRPLAALAQRIFTASLPPEKNCAGADHICQDNLGS	784
Qy	785	FSPMSLDCLVVGGRPREPNVTVVRNDCEDSVRTQVTEFFPLDLSYRKVSTLQNRORSW	844
Dd	785	FSPFGLKSLVGSNLELNAEVWMWNDCEDSIGTTITFSHPAGUSYRYVABEQQOGQURSL	844
Qy	845	RLACESASTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTS	904
Dd	845	HLTC--CSAPVSGQWTSTSCRINHLIFRGAQITFLATFDVPKAVGLDRLLLLIANVSS	902
Qy	905	ENNMPRNKTETEFOLELPKVYAVVMVTSHGVS TKYNLFNTAS-ENTSRVMOHQYQVSNLQG	963
Dd	903	ENNIPRTSKTIFOLELPKVYAVYVSSHEQTKYLNFSSEBEKESHVAMHYOVNNLQG	962
Qy	964	RSULPISLVFLPVRLNQTVIWDPRQVTFSENLSSTCHTKERLDSHSDFLAELKAPVUNC	1023
Dd	963	RDLFVSNFNPVVELNOQEAVMMDVEVSHPNPSLRCSCEAIAPASPDLAHIOKNPVLDC	1022
Qy	1024	SIACVORIQCIDIFFGIQEEFNATLKGNLDFWYIKTSHNHLLIVSTAELFNDVSFTLL	1083
Dd	1023	SIACLFRCDVPSFSVQEBELDTLKGNLSEFWVRQILOKKVSVSVSAEIIFTSVYSQL	1082
Qy	1084	PGOGAFVRSOTETKVEPFPENPLPIVSGSVGLLLLALITAALYKLGFPFKQYKDMMS	1143
Dd	1083	PGOEFNRAQTIITVLEKYKHNPPIPLIVGSSIGGLLLALLAJITAVLYKVGFQEKYEMME	1142
Qy	1144	E 1144	
Dd	1143	E 1143	

RESULT 14

RESULT 14
US-09-350-259-2

US-09-350-259-2 : Sequence 2: Application IIS/09350259

; sequence 2, APPLICATION US
: Patent No. US20020062008A1

; FALCEN NO. US2002006
: GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.

APPLICANT: GARRATT, MICHAEL W.
; APPLICANT: Van der Vieren, Monica

Qy 725 DPVSPVLRNLSVLTGTPLSAFGNLRPVLAEADAORLFTALFPFPEKNCNDNI CODDLSIT 784
Db 722 DVSPPIILHLNLSLVREPIPSQNLRLPVLAVGSDQLFTASLPFPEKNCQDGLCEGLGVT 781
Qy 785 FSNMSLCLVVGGPREFNVTVVRNDEGDSYRTQVTFPPFLDLRYKRVSTLQNRORSW 844
Db 782 LSFSGLOTLTVGSSLELNVIIVVMNAGEDSVTVVSLYPAGLSHRRVSGNQKOPHOSAL 841
Qy 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
Db 842 RLACETV-PTDEG-LRSSRCSVNHPIFHEGSGNCTFIPTDVSYSKATLGDRLMRASASS 899
Qy 905 ENNMPRTNKTFFOLELPVKYAVVMVTVSHGVSTKYLPF-TASENTRVMQHOYQVSNLQ 963
Db 900 ENNKASSSKATFQLELPVKYAVVTMISROEESTKYFNPFATSDKEMKEAHEHYRVNLSQ 959
Qy 964 RSLPISLVLPVRLNQTIVWDRPQVTFPSENLSSTCHTKERLPSHSDPLAELRKAPVVC 1023
Db 960 RDLAISINFWPVLNGVAVMDVMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDC 1017
Qy 1024 STAVCQRTQCDIPFGIQQEENATLKNLSFDWYIKTSHNHLIVSTABILLFNDSVFTLL 1083
Db 1018 SIADCLQPCDVPSPSQEELDFTLKGNLSFGWVRETLQKKVLVVSAEITFTDTSVSQL 1077
Qy 1084 PQGAFVRSQETKVEPEVNPPLIIVGSSVGLLALLALITAAIYKLGFFKRYKDMMS 1143
Db 1078 PQEAFMRAQMEMVLEDEVYNAIPIIMGSSVGALLLALLITATLYKLGFFKRYKEMLE 1137
Qy 1144 E 1144
Db 1138 D 1138

RESULT 15

US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US2003007278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 58.0%; Score 3455; DB 11; Length 1161;
Best Local Similarity 59.9%; Pred. No. 2.9e-308;
Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;

Qy 5 VLLLTALTLCGFNLDENAMTFQENARGFQSVQLOGSRVVGAPQBIIVANQSGSLY 64
Db 5 VLLLSVLAHYGFNLDVEEPTFIQEDAGGFQSVQFQGSRLVVGAPLEVVAANTGRLY 64
Qy 65 QCDYSTGSGCEPIRLQVPEAVNMSGLSLAATSPQLLIACGPTVHQTCSNTYKGLCF 124
Db 65 DCAATGMCQPIPLHTRPEAVNMSGLTLAASTNGSRLLIACGPTLHVRVCGENSYKGSCL 124

Qy 125 LFGSNLRQOPKPEALRGCPQEDSDIAFLIDSGSIIIPHFRMRKEFVSTVMEQLKSK 184
Db 125 LLSRW-ELIQTVPDAPCEPHQEMDI VFLIDSGSIDQNDPNQMKGFQVAVMQEGEGTD 183
Qy 185 TLFSLMOYSEEFRIHFTFEFQNNPNRSLVKPITOLLGRTHTATGIRKVVRELFINITNG 244
Db 184 TLFALMOYSNLLKHFTFTQFRTSPSQSLVDPIVQLKGLTFTATGILTVVTLQFLHKG 243
Qy 245 ARKNAKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPRSEKSKQELNTIA 304
Db 244 ARKSAKKILVITDQKYKDPLEYSVIVQAEKAGIIRIYAIGVGHAFQGTARQELNTIS 303
Qy 305 SKPEPRHVFVNNFEALKTIQNLQREKIFAIEGTQTCSSSFEHEMSQEGFSAAITNSGP 364
Db 304 SAPQDHPVKVDNFAALGSIQKQLEKIYAVEGTQSRASSSPQHEMSQEGFSTALTMGDL 363
Qy 365 LLSTVSGSYWAGVFLYTSKEKSTFINMTRVDSMDMDAYLCYAAAIIILNRVQSILVLAGP 424
Db 364 FLGAVGSFSWGGAFLYPPNMSPTFINMSQENYMDRDSYLGSTELALMKGVQNLVLAGP 423
Qy 425 RYOHIGLVAMFRONTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDLVLGAPHYE 484
Db 424 RYOHIGKAVITQVSRQWRKKAETGTQIGSYFGASLCSVDVDSNGSTDLVLGAPHYE 483
Qy 485 QTRGGQVSVCLPRGQRAARQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDAVIGAP 544
Db 484 QTRGGQVSVCLPRGQRAARQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDAVIGAP 543
Qy 545 GEBDNRAVLYFHTSGSISPSHSQRIAGSKLSPRLOYFGOSLSSGQDGLTMDGLVLT 604
Db 544 GEQENRAVLYFHTSGSISPSHSQRIAGSKLSPRLOYFGOSLSSGQDGLTMDGLVLT 603
Qy 605 GAQGHVLLRSQPLRVKATMEFNPREVARNVFNCDQVVKGEAGVRVCHLVOKSTRD 664
Db 604 GARGQVLLRSQPLRVKATMEFNPREVARNVFNCDQVVKGEAGVRVCHLVOKSTRD 663
Qy 665 RLREGQIQSVVTVDLALDLSGRPHSRVFNETQNSTRTQTVGLGTOTCETLKLQLPNCIE 724
Db 664 QL--GDIOSSVRFDLALDPLGRLTSRAIFNETQNTPLTRRKTGLGHCETLKLQLPNCIE 721
Qy 725 DPVSPVLRNLSVLTGTPLSAFGNLRPVLAEADAORLFTALFPFPEKNCNDNI CODDLSIT 784
Db 722 DVSPPIILHLNLSLVREPIPSQNLRLPVLAVGSDQLFTASLPFPEKNCQDGLCEGLGVT 781
Qy 785 FSNMSLCLVVGGPREFNVTVVRNDEGDSYRTQVTFPPFLDLRYKRVSTLQNRORSW 844
Db 782 LSFSGLOTLTVGSSLELNVIIVVMNAGEDSVTVVSLYPAGLSHRRVSGNQKOPHOSAL 841
Qy 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
Db 842 RLACETV-PTDEG-LRSSRCSVNHPIFHEGSGNCTFIPTDVSYSKATLGDRLMRASASS 899
Qy 905 ENNMPRTNKTFFOLELPVKYAVVMVTVSHGVSTKYLPF-TASENTRVMQHOYQVSNLQ 963
Db 900 ENNKASSSKATFQLELPVKYAVVTMISROEESTKYFNPFATSDKEMKEAHEHYRVNLSQ 959
Qy 964 RSLPISLVLPVRLNQTIVWDRPQVTFPSENLSSTCHTKERLPSHSDPLAELRKAPVVC 1023
Db 960 RDLAISINFWPVLNGVAVMDVMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDC 1017
Qy 1024 STAVCQRTQCDIPFGIQQEENATLKNLSFDWYIKTSHNHLIVSTABILLFNDSVFTLL 1083
Db 1018 SIADCLQPCDVPSPSQEELDFTLKGNLSFGWVRETLQKKVLVVSAEITFTDTSVSQL 1077
Qy 1084 PQGAFVRSQETKVEPEVNPPLIIVGSSVGLLALLALITAAIYKLGFFKRYKDMMS 1143
Db 1078 PQEAFMRAQMEMVLEDEVYNAIPIIMGSSVGALLLALLITATLYKLGFFKRYKEMLE 1137
Qy 1144 E 1144
Db 1138 D 1138

Tue Nov 25 14:51:51 2003

us-09-902-481b-1.rapb

Page 13

Search completed: November 25, 2003, 14:41:06
Job time: 33.4941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:13:09 ; Search time 13.7527 Seconds
(without alignments)
3547.268 Million cell updates/sec

Title: US-09-902-481B-1

Perfect score: 5956

Sequence: 1 MALRVLLLTALTLCGHNLD.....FKROYKDMSEGGPPGAPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5956	100.0	1153	1	US-08-173-497-3
2	5956	100.0	1153	1	US-08-286-889-3
3	5956	100.0	1153	1	US-08-485-618-3
4	5956	100.0	1153	1	US-08-362-652-3
5	5956	100.0	1153	2	US-08-605-672-3
6	5956	100.0	1153	2	US-08-482-293A-3
7	5956	100.0	1153	2	US-08-943-363-3
8	5956	100.0	1153	3	US-09-193-043-3
9	5956	100.0	1153	4	US-09-688-307A-3
10	5925.5	99.5	1152	2	US-08-476-062A-43
11	5925.5	99.5	1152	5	PCT-US96-01314-43
12	5925.5	99.5	1152	6	Patent No. 5424399-2
13	3503	58.8	1163	2	US-08-476-062A-44
14	3503	58.8	1163	5	PCT-US96-01314-44
15	3480	58.4	1163	1	US-08-173-497-4
16	3480	58.4	1163	1	US-08-286-889-4
17	3480	58.4	1163	1	US-08-485-618-4
18	3480	58.4	1163	2	US-08-362-652-4
19	3480	58.4	1163	2	US-08-605-672-4
20	3480	58.4	1163	2	US-08-482-293A-4
21	3480	58.4	1163	2	US-08-943-363-4
22	3480	58.4	1163	3	US-09-193-043-4
23	3480	58.4	1163	4	US-09-688-307A-4
24	3455	58.0	1161	1	US-08-173-497-2
25	3455	58.0	1161	1	US-08-286-889-2
26	3455	58.0	1161	1	US-08-485-618-2
27	3455	58.0	1161	1	US-08-362-652-2

28 3455 58.0 1161 2 US-08-605-672-2 Sequence 2, Appli
29 3455 58.0 1161 2 US-08-482-293A-2 Sequence 2, Appli
30 3455 58.0 1161 2 US-08-943-363-2 Sequence 2, Appli
31 3455 58.0 1161 3 US-09-193-043-2 Sequence 2, Appli
32 3455 58.0 1161 4 US-09-688-307A-2 Sequence 2, Appli
33 3439.5 57.7 1161 1 US-08-485-618-99 Sequence 99, Appli
34 3439.5 57.7 1161 2 US-08-605-672-99 Sequence 99, Appli
35 3439.5 57.7 1161 2 US-08-482-293A-99 Sequence 99, Appli
36 3439.5 57.7 1161 2 US-08-943-363-99 Sequence 99, Appli
37 3439.5 57.7 1161 3 US-09-193-043-99 Sequence 99, Appli
38 3439.5 57.7 1161 4 US-09-688-307A-99 Sequence 99, Appli
39 3270.5 54.9 1161 3 US-09-193-043-55 Sequence 55, Appli
40 3270.5 54.9 1161 4 US-09-688-307A-55 Sequence 55, Appli
41 3264 54.8 1161 1 US-08-485-618-53 Sequence 53, Appli
42 3264 54.8 1161 1 US-08-362-652-53 Sequence 53, Appli
43 3264 54.8 1161 2 US-08-605-672-53 Sequence 53, Appli
44 3264 54.8 1161 2 US-08-482-293A-53 Sequence 53, Appli
45 3264 54.8 1161 2 US-08-943-363-53 Sequence 53, Appli

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/173.497
; APPLICATION NUMBER: US/08/173.497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFQGSVVQLGSRVVVGAPOEIVAAQR 60
DB 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFQGSVVQLGSRVVVGAPOEIVAAQR 60

QY 61 GSYQCDYSTGSCPEIRLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
DB 61 GSYQCDYSTGSCPEIRLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
QY 121 GLCLFGLSGLNRQOQPKPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEFYSTVMEQL 180
DB 121 GLCLFGLSGLNRQOQPKPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEFYSTVMEQL 180
QY 181 KKSXTLSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KKSXTLSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQEL 300
DB 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQEL 300
QY 301 NTIAKPPRDHVQVNFALKTIONQLREKIPAEIGTQTGSSSSFEHMSQEGFSAIT 360
DB 301 NTIAKPPRDHVQVNFALKTIONQLREKIPAEIGTQTGSSSSFEHMSQEGFSAIT 360
QY 361 SNGPLSTVGSYDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLV 420
DB 361 SNGPLSTVGSYDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLV 420
QY 421 LGAPRYOHIGLVANFRQNTGWNESNANVKGTOICAYFGASLCSVDVDSNGSTDLDVLIGAP 480
DB 421 LGAPRYOHIGLVANFRQNTGWNESNANVKGTOICAYFGASLCSVDVDSNGSTDLDVLIGAP 480
QY 481 HYYEQTRGGQVSVCPLRGRQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEQTRGGQVSVCPLRGRQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
QY 541 IGARCEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDITMDGLV 600
DB 541 IGARCEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDITMDGLV 600
QY 601 DLTGVAQGHVLLLSQVLRVKAIMEFNPREVARNVPECNDQVVKGEAGVRVCLHVQK 660
DB 601 DLTGVAQGHVLLLSQVLRVKAIMEFNPREVARNVPECNDQVVKGEAGVRVCLHVQK 660
QY 661 STRDLREGQIQSVVYTDALDSGRPHSRVAFNETKSTRQTVGLGTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVYTDALDSGRPHSRVAFNETKSTRQTVGLGTQTCETLKLQLP 720
QY 721 NCIEDPVSPVLRNLSLVGTPLSAFGLNRPVLAEDAQRLLTALPFPEKNCNGDNIQDD 780
DB 721 NCIEDPVSPVLRNLSLVGTPLSAFGLNRPVLAEDAQRLLTALPFPEKNCNGDNIQDD 780
QY 781 LSITFSFMSLCLVVGGRPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
DB 781 LSITFSFMSLCLVVGGRPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
QY 841 QRSWLACASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENNPRNTKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMQYQVSN 960
DB 901 NVTSENNPRNTKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMQYQVSN 960
QY 961 LGORSLPLSLVFLVPVRNQTVMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPV 1020
DB 961 LGORSLPLSLVFLVPVRNQTVMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPV 1020
QY 1021 VNCSTAVCQRIQCDIPFGIQQEENFATLKGNSLFDWYIKTSHNHLIIIVSTAEIIFNDSVF 1080
DB 1021 VNCSTAVCQRIQCDIPFGIQQEENFATLKGNSLFDWYIKTSHNHLIIIVSTAEIIFNDSVF 1080
QY 1081 TLLPQCGAFVRSQTTKTEPEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKQRYKD 1140
DB 1081 TLLPQCGAFVRSQTTKTEPEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKQRYKD 1140

QY 1141 MMSEGPPGAEPO 1153
DB 1141 MMSEGPPGAEPO 1153

RESULT 2

US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-286-889-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCCHGNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVANQR 60
DB 1 MALRVLLLTALTLCCHGNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVANQR 60
QY 61 GSYQCDYSTGSCPEIRLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
DB 61 GSYQCDYSTGSCPEIRLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVK 120
QY 121 GLCLFGLSGLNRQOQPKPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEFYSTVMEQL 180
DB 121 GLCLFGLSGLNRQOQPKPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEFYSTVMEQL 180
QY 181 KKSXTLSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KKSXTLSLMQYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQEL 300
DB 241 ITNGARKNAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQEL 300


```
QY 301 NTIASKPPRDHVPQVNNFEALKTIONLREKIFAIBGTQTGSSSSFEHMSQBGFSAAIT 360
Db 301 NTIASKPPRDHVPQVNNFEALKTIONLREKIFAIBGTQTGSSSSFEHMSQBGFSAAIT 360
QY 361 SNGPLLSYVGSYDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAIILNRVQSLV 420
Db 361 SNGPLLSYVGSYDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFQNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAP 480
Db 421 LGAPRYQHIGLVAMFQNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAP 480
QY 481 HYIEQTRGGQVSCVPLPRQARAWQCDVLYGEGQGWGRFGAALTVLGVDVNGDKLTDVA 540
Db 481 HYIEQTRGGQVSCVPLPRQARAWQCDVLYGEGQGWGRFGAALTVLGVDVNGDKLTDVA 540
QY 541 IGAPGEDNRGAVYLHGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTWDGLV 600
Db 541 IGAPGEDNRGAVYLHGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTWDGLV 600
QY 601 DLTVGAQGHVLLRSOPVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVVRCLHVQK 660
Db 601 DLTVGAQGHVLLRSOPVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVVRCLHVQK 660
QY 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAFVNETKNSTRRTQVGLTQTCTETKLQLP 720
Db 661 STRDLREGQIQSVVTVYDLALDSGRPHSRVAFVNETKNSTRRTQVGLTQTCTETKLQLP 720
QY 721 NCIEDPVSPIVLRNLSLVTPLSAGNLRPLVLAEDAQRFTALFPFKNCNGNDNICODD 780
Db 721 NCIEDPVSPIVLRNLSLVTPLSAGNLRPLVLAEDAQRFTALFPFKNCNGNDNICODD 780
QY 781 LSITFSFMSLDCLVCGPREFNTVTVRNDGDSYRTQVTFPPPLDLVSRKVSSTLQNRS 840
Db 781 LSITFSFMSLDCLVCGPREFNTVTVRNDGDSYRTQVTFPPPLDLVSRKVSSTLQNRS 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGKLLKA 900
Db 841 QRSWLACESASSTEVSGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGKLLKA 900
QY 901 NVTSENNMPTNTEFOLELPVKYAVYVWVTSYGVSTKYLNFTASENSTRVMOHQVSN 960
Db 901 NVTSENNMPTNTEFOLELPVKYAVYVWVTSYGVSTKYLNFTASENSTRVMOHQVSN 960
QY 961 LGORSPLISLVLPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDPLAELKAPV 1020
Db 961 LGORSPLISLVLPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDPLAELKAPV 1020
QY 1021 VNCISIAVCQRIQCDIPFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSF 1080
Db 1021 VNCISIAVCQRIQCDIPFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSF 1080
QY 1081 TLLPGQGAFAVRSOTETKVEFEPVNPPLIVGSSVGLLALLITAAIYKLGFFKQYKD 1140
Db 1081 TLLPGQGAFAVRSOTETKVEFEPVNPPLIVGSSVGLLALLITAAIYKLGFFKQYKD 1140
QY 1141 MMSEGGPPGAEPQ 1153
Db 1141 MMSEGGPPGAEPQ 1153
```

RESULT 3

```
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
```

```
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-3
```

```
Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGHNLDNTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANOR 60
Db 1 MALRVLLLTALTLCGHNLDNTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANOR 60
QY 61 GSYLYQCDYSTGSCPEIRLQVPVEAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
Db 61 GSYLYQCDYSTGSCPEIRLQVPVEAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
QY 121 GLCFLPGSNLRQOPQKFPPEALRCCQEDSDIAFLIDGSGSIIPHDPFRMKFYSTVMEQL 180
Db 121 GLCFLPGSNLRQOPQKFPPEALRCCQEDSDIAFLIDGSGSIIPHDPFRMKFYSTVMEQL 180
QY 181 KKSKTILFSLMOYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KKSKTILFSLMOYSEEFRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAFRSEKSRQEL 300
QY 301 NTIASKPPRDHVPQVNNFEALKTIONLREKIFAIBGTQTGSSSSFEHMSQBGFSAAIT 360
Db 301 NTIASKPPRDHVPQVNNFEALKTIONLREKIFAIBGTQTGSSSSFEHMSQBGFSAAIT 360
QY 361 SNGPLLSYVGSYDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAIILNRVQSLV 420
Db 361 SNGPLLSYVGSYDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFQNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAP 480
```

Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
Qy 481 HYYEOTRGQVSCVPLPRQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYYEOTRGQVSCVPLPRQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPCEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQOSLGGGDLTMDGLV 600
Db 541 IGAPCEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQOSLGGGDLTMDGLV 600
Qy 601 DLTGAGQHVLLRSQVPLRVKAIMFNPVFECDNVKKEAGEVRCVCHVOK 660
Db 601 DLTGAGQHVLLRSQVPLRVKAIMFNPVFECDNVKKEAGEVRCVCHVOK 660
Qy 661 STRDLRREGQISQVVTYDIALDSGRPHSAVNETKNSRRTQVGLTQTCTETIKLQLP 720
Db 661 STRDLRREGQISQVVTYDIALDSGRPHSAVNETKNSRRTQVGLTQTCTETIKLQLP 720
Qy 721 NCIEDPVSIVLRNPLSVGTPLSAFNGLRPVLAQAQLFTALPPFKNCNDNICODD 780
Db 721 NCIEDPVSIVLRNPLSVGTPLSAFNGLRPVLAQAQLFTALPPFKNCNDNICODD 780
Qy 781 LSTTFEFSMLDCLVCGPREFNVTVVRNDGDSYRTQVTFEPLDLSYRKUSTLONORS 840
Db 781 LSTTFEFSMLDCLVCGPREFNVTVVRNDGDSYRTQVTFEPLDLSYRKUSTLONORS 840
Qy 841 QRSWRLACASASTSVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWRLACASASTSVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNMPTNKTEFQLELPLVKYAVYVMVTSKGVSSTKYLNFTASENTSRVNHQYQVSN 960
Db 901 NVTSENNMPTNKTEFQLELPLVKYAVYVMVTSKGVSSTKYLNFTASENTSRVNHQYQVSN 960
Qy 961 LQORSUPLISLVFLVPLRNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Db 961 LQORSUPLISLVFLVPLRNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
Qy 1021 VNCISAVCORIOCDIPFGIQEFNATLKNLSFDMYIKTSHNHLIIVSTABILFNDVSF 1080
Db 1021 VNCISAVCORIOCDIPFGIQEFNATLKNLSFDMYIKTSHNHLIIVSTABILFNDVSF 1080
Qy 1081 TLLPGGAFVRSQTEKVPFEPVNPPLPLIVGSSVGLLALLITAALYKLGFFKRYQKD 1140
Db 1081 TLLPGGAFVRSQTEKVPFEPVNPPLPLIVGSSVGLLALLITAALYKLGFFKRYQKD 1140
Qy 1141 MMSEGGPPGAEPQ 1153
Db 1141 MMSEGGPPGAEPQ 1153

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362.652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173.497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286.889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTLCGHNLDFTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAAQR 60
Db 1 MALRVLLLTALTLCGHNLDFTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAAQR 60
Qy 61 GSLVQCDYSTCSCSPILRQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTVVK 120
Db 61 GSLVQCDYSTCSCSPILRQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTVVK 120
Qy 121 GLCLFGLNSLRQQKQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMBQL 180
Db 121 GLCLFGLNSLRQQKQKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMBQL 180
Qy 181 KKSCTLFSLMYSSEFRIFHTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
Db 181 KKSCTLFSLMYSSEFRIFHTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQEL 300
Qy 301 NTIASKPPRDHVFOVNNFEALKTIONQLREKIFAIEGTQTGSSSSSFEHMSQEGFSAAIT 360
Db 301 NTIASKPPRDHVFOVNNFEALKTIONQLREKIFAIEGTQTGSSSSSFEHMSQEGFSAAIT 360
Qy 361 SNGPLLSVSGYDWMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Db 361 SNGPLLSVSGYDWMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
Qy 481 HYYEOTRGQVSCVPLPRQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Db 481 HYYEOTRGQVSCVPLPRQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
Qy 541 IGAPCEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQOSLGGGDLTMDGLV 600
Db 541 IGAPCEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQOSLGGGDLTMDGLV 600
Qy 601 DLTGAGQHVLLRSQVPLRVKAIMFNPVFECDNVKKEAGEVRCVCHVOK 660
Db 601 DLTGAGQHVLLRSQVPLRVKAIMFNPVFECDNVKKEAGEVRCVCHVOK 660

601 DLTGAGQHVLRLSRQVLRVKAIBFNPENPREVARNVFNCDQVVKKEAGEVRVCLHVOK 660
661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKNSRRQTQVGLTQTCTELKQLP 720
661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKNSRRQTQVGLTQTCTELKQLP 720
721 NCIEDPVPVILRLNLSVGTPLSAFNGNLRPVLAEDAQRLLFTALFFPKNCNDNICQDD 780
721 NCIEDPVPVILRLNLSVGTPLSAFNGNLRPVLAEDAQRLLFTALFFPKNCNDNICQDD 780
781 LSITFSFMSLDCVLVGGPREFNTVTVRNDGDSYRTQVTFPPDLDSYRKVSTLQNR 840
781 LSITFSFMSLDCVLVGGPREFNTVTVRNDGDSYRTQVTFPPDLDSYRKVSTLQNR 840
841 QRSWLACASASSTEVSGALKSTSCSINHPIFFENSEVFNITFDVDSKASLGNKLLKA 900
841 QRSWLACASASSTEVSGALKSTSCSINHPIFFENSEVFNITFDVDSKASLGNKLLKA 900
901 NVTSENMPRTNKTETQLELPLVYAVYVMTVSHGVSTKYNFTASNTSRVQHQVQSN 960
901 NVTSENMPRTNKTETQLELPLVYAVYVMTVSHGVSTKYNFTASNTSRVQHQVQSN 960
961 LGORSPLISLVLVPLVRLNQTIVMDRPOVTFESENLSSTCHTKERLPSHDSFLAELRKAPV 1020
961 LGORSPLISLVLVPLVRLNQTIVMDRPOVTFESENLSSTCHTKERLPSHDSFLAELRKAPV 1020
1021 VNCISIAVCQRIQCDIPFFGQIEBFNATLKGNSLFDWYIKTSHNHLIVSTAELLFNDVSF 1080
1021 VNCISIAVCQRIQCDIPFFGQIEBFNATLKGNSLFDWYIKTSHNHLIVSTAELLFNDVSF 1080
1081 TLLPGGAFVRSOTETKVPFVNPPLIVGVSSVGLLALLITAAALKYKGFKQYKD 1140
1081 TLLPGGAFVRSOTETKVPFVNPPLIVGVSSVGLLALLITAAALKYKGFKQYKD 1140
1141 MMSEGGPPGAEPQ 1153
1141 MMSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
Sequence 3, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994

ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-3
Query Match 100.0%; Score 5956; DB 2; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEI1VAANOR 60
DB 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEI1VAANOR 60
QY 61 GSLYQCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSSENTYVK 120
DB 61 GSLYQCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSSENTYVK 120
QY 121 GLCFLFGSNLRQOPKQPEALRGCPQEDSDIAFLIDGSGSII1PHDFRMKEFVSTVMEOL 180
DB 121 GLCFLFGSNLRQOPKQPEALRGCPQEDSDIAFLIDGSGSII1PHDFRMKEFVSTVMEOL 180
QY 181 KSKTFLSMLQYSEEFHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
DB 181 KSKTFLSMLQYSEEFHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFSEKSRQEL 300
QY 301 NTIASKPPRDHVFQNNFEALKTIONQREKIFAIEGTQTGSSSSSFEHMSQEGFSAAIT 360
DB 301 NTIASKPPRDHVFQNNFEALKTIONQREKIFAIEGTQTGSSSSSFEHMSQEGFSAAIT 360
QY 361 SNGPLLSVTVGSDYDAGGVFLYTSKSTFINMTRVDSMDNDAYLGVAAAIILNRNVQSLV 420
DB 361 SNGPLLSVTVGSDYDAGGVFLYTSKSTFINMTRVDSMDNDAYLGVAAAIILNRNVQSLV 420
QY 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
QY 481 HYYEOTRGQSVSVCPPLRGQARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEOTRGQSVSVCPPLRGQARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVA 540
QY 541 ICAPGEEDNRGAVLPHGTSGSGISPSHSORTAGSKLSPRLOYFGOSLGGGDLTMDGLV 600
DB 541 ICAPGEEDNRGAVLPHGTSGSGISPSHSORTAGSKLSPRLOYFGOSLGGGDLTMDGLV 600
QY 601 DLTGAGQHVLRLSRQVLRVKAIMEFNPENPREVARNVFNCDQVVKKEAGEVRVCLHVOK 660
DB 601 DLTGAGQHVLRLSRQVLRVKAIMEFNPENPREVARNVFNCDQVVKKEAGEVRVCLHVOK 660
QY 661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKNSRRQTQVGLTQTCTELKQLP 720
DB 661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKNSRRQTQVGLTQTCTELKQLP 720
QY 721 NCIEDPVPVILRLNLSVGTPLSAFNGNLRPVLAEDAQRLLFTALFFPKNCNDNICQDD 780
DB 721 NCIEDPVPVILRLNLSVGTPLSAFNGNLRPVLAEDAQRLLFTALFFPKNCNDNICQDD 780

Qy 781 LSITFSMSLDCVLVGGPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
Db 781 LSITFSMSLDCVLVGGPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
Qy 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
Db 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNPRNTKTEFOLELVKAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
Db 901 NVTSENNPRNTKTEFOLELVKAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
Qy 961 LQSRSLPISLVFLVPRLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPV 1020
Db 961 LQSRSLPISLVFLVPRLNQTVIWDPRQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPV 1020
Qy 1021 VNCISAVCQRIQCDDIPFGIOEFENATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDVSF 1080
Db 1021 VNCISAVCQRIQCDDIPFGIOEFENATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDVSF 1080
Qy 1081 TLLPQGGAFVRSQTEKVEPEVPNPLPLIVGSSVGGLLALITAALYKLGFFKROYKD 1140
Db 1081 TLLPQGGAFVRSQTEKVEPEVPNPLPLIVGSSVGGLLALITAALYKLGFFKROYKD 1140
Qy 1141 MMSEGGPPGABPQ 1153
Db 1141 MMSEGGPPGABPQ 1153

RESULT 6
US-08-482-293A-3
Sequence 3, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 381659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-3
Query Match 100.0%; Score 5956; DB 2; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRVLLLTALTICHGFNLDTENAMTFQENARFGQSVVQLQSRVVVGAPOEIVANQR 60
Db 1 MALRVLLLTALTICHGFNLDTENAMTFQENARFGQSVVQLQSRVVVGAPOEIVANQR 60
Qy 61 GSIYQCDYSYSGSCEPTRLQVPVEAVNMSLGLSLAATSPPOLACAGTPTVHQTSENTRYK 120
Db 61 GSIYQCDYSYSGSCEPTRLQVPVEAVNMSLGLSLAATSPPOLACAGTPTVHQTSENTRYK 120
Qy 121 GLCFLFGSNLRQQPKFPPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEFVSTVMEQL 180
Db 121 GLCFLFGSNLRQQPKFPPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEFVSTVMEQL 180
Qy 181 KSKTLPFLSMQYSEEPRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Db 181 KSKTLPFLSMQYSEEPRIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQEL 300
Qy 301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIIPAIECTQTGSSSSFEHMSQEGFSAIT 360
Db 301 NTIASKPPRDHVFQVNNFEALKTIONQLREKIIPAIECTQTGSSSSFEHMSQEGFSAIT 360
Qy 361 SNGPLLTSGSYDWAGVFLYTSKEKSTFNTMTVDSMDNDAYLGAAAAIILNRVQSLV 420
Db 361 SNGPLLTSGSYDWAGVFLYTSKEKSTFNTMTVDSMDNDAYLGAAAAIILNRVQSLV 420
Qy 421 LGAPRYOHIGLVAMFRONTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAP 480
Db 421 LGAPRYOHIGLVAMFRONTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAP 480
Qy 481 HYYEQTRGGQVSVCPFLPRGQARQWQCDVLYGEGQPMGRFGAALTVLGDNVGDKLTDVA 540
Db 481 HYYEQTRGGQVSVCPFLPRGQARQWQCDVLYGEGQPMGRFGAALTVLGDNVGDKLTDVA 540
Qy 541 IGAPGEDNARGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDGLTMDGLV 600
Db 541 IGAPGEDNARGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDGLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGVRVCLHVQK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAGVRVCLHVQK 660
Qy 661 STRDLREGQIQSVVTVYDLALDSGRPHSRAVFNETKSTRQTOVLGLTQTCETLKLQLP 720
Db 661 STRDLREGQIQSVVTVYDLALDSGRPHSRAVFNETKSTRQTOVLGLTQTCETLKLQLP 720
Qy 721 NCIEDPVSPVLRNLSLVGTPLSAFNLRPVLAEDAQRLETPALPFPEKNCNDNICDD 780
Db 721 NCIEDPVSPVLRNLSLVGTPLSAFNLRPVLAEDAQRLETPALPFPEKNCNDNICDD 780
Qy 781 LSITFSMSLDCVLVGGPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
Db 781 LSITFSMSLDCVLVGGPRENVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNR 840
Qy 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
Db 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENNPRNTKTEFOLELVKAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 960

Db 901 NVTSENMPRTNKTEFQLELPVKYAVYVMTVSHGVSTKYNFTASENTSRVMQHOYQVSN 960
Qy 961 LGQSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPV 1020
Db 961 LGQSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPV 1020
Qy 1021 VNCISIAVCQRIQCDIPFFGIQBEFNATLKGNSLFDWYIKTSNHNLLIVSTABILFNDVSF 1080
Db 1021 VNCISIAVCQRIQCDIPFFGIQBEFNATLKGNSLFDWYIKTSNHNLLIVSTABILFNDVSF 1080
Qy 1081 TLLPGGAFVRSOTETKVPFEPVNPPLIVGVSSVGGLLLLALITAAALYKLGFFKQYKD 1140
Db 1081 TLLPGGAFVRSOTETKVPFEPVNPPLIVGVSSVGGLLLLALITAAALYKLGFFKQYKD 1140
Qy 1141 MMSEGGPPGAEPQ 1153
Db 1141 MMSEGGPPGAEPQ 1153

RESULT 7

US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-943-363-3

Query Match 100.0%; Score 5956; DB 2; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLTALTLCCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEI VAAQNR 60
Db 1 MALRVLLTALTLCCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEI VAAQNR 60
Qy 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATTSPPQLLAGCPTVHTCSENTYVK 120
Db 61 GSLYQCDYSTGCEPIRLQVPVEAVNMSLGLSLAATTSPPQLLAGCPTVHTCSENTYVK 120
Qy 121 GLCFLFGSNLRQPOKPFPEALRGCPQEDSDTAFLIDGSGSIIPHDFRMRKEFVSTVMBQL 180
Db 121 GLCFLFGSNLRQPOKPFPEALRGCPQEDSDTAFLIDGSGSIIPHDFRMRKEFVSTVMBQL 180
Qy 181 KKSCTLFSLMQYSEBFRIFHTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
Db 181 KKSCTLFSLMQYSEBFRIFHTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAPRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVITPEADREGVIRYVIGVDAPRSEKSRQEL 300
Qy 301 NTIASKPPRDHVFQVNNFEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAT 360
Db 301 NTIASKPPRDHVFQVNNFEALKTIQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRNVQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRNVQSLV 420
Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYTEOTRGQSVSVCPPLRGQARWOCDAVLGEGQCPWGRFGCAALTVLGDVNGDKLTQVA 540
Db 481 HYTEOTRGQSVSVCPPLRGQARWOCDAVLGEGQCPWGRFGCAALTVLGDVNGDKLTQVA 540
Qy 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Db 541 IGAPGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDNDVVKGKEAGEVRVCLHVOK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDNDVVKGKEAGEVRVCLHVOK 660
Qy 661 STRDRLRGQIQSVVTVYDLALDSGRPHSRVFNENKSTRROTQVLGLTQTCETLKLQUP 720
Db 661 STRDRLRGQIQSVVTVYDLALDSGRPHSRVFNENKSTRROTQVLGLTQTCETLKLQUP 720
Qy 721 NCIEDPVSPIVLRLNFSLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFFKKNCGNDNICQDD 780
Db 721 NCIEDPVSPIVLRLNFSLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFFKKNCGNDNICQDD 780
Qy 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDSYRTQVTFPPFLDLSYRKVSTLQNR 840
Db 781 LSITFSFMSLDCLVVGGPREFNVTVVRNDSYRTQVTFPPFLDLSYRKVSTLQNR 840
Qy 841 QRSWELACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWELACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSENMPRTNKTEFQLELPVKYAVYVMTVSHGVSTKYNFTASENTSRVMQHOYQVSN 960
Db 901 NVTSENMPRTNKTEFQLELPVKYAVYVMTVSHGVSTKYNFTASENTSRVMQHOYQVSN 960
Qy 961 LGQSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPV 1020
Db 961 LGQSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPV 1020
Qy 1021 VNCISIAVCQRIQCDIPFFGIQBEFNATLKGNSLFDWYIKTSNHNLLIVSTABILFNDVSF 1080
Db 1021 VNCISIAVCQRIQCDIPFFGIQBEFNATLKGNSLFDWYIKTSNHNLLIVSTABILFNDVSF 1080

SEQ ID NO 3
LENGTH: 1153
TYPE: PRT
ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 100.0%; Score 5956; DB 4; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCCHGFNDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVANQR 60
DB 1 MALRVLLLTALTLCCHGFNDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVANQR 60
QY 61 GSYLQCDYGTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
DB 61 GSYLQCDYGTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120
QY 121 GLCFLFGSNLRQOPKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 180
DB 121 GLCFLFGSNLRQOPKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL 180
QY 181 KSKTLFSLMQYSEEPRIHFTKPEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
DB 181 KSKTLFSLMQYSEEPRIHFTKPEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELPN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYGVGDAFSEKSRQEL 300
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYGVGDAFSEKSRQEL 300
QY 301 NTIASKPPRDHVFQNNFALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAIT 360
DB 301 NTIASKPPRDHVFQNNFALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAIT 360
QY 361 SNGPLSTVGSVDWAGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
DB 361 SNGPLSTVGSVDWAGVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLV 420
QY 421 LGAPRYOHLGLVAMPONTGMESNANVGTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
DB 421 LGAPRYOHLGLVAMPONTGMESNANVGTQIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
QY 481 HYYEOTRGQSVCPPLRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEOTRGQSVCPPLRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
QY 541 IGAPGEDNRGAVYLPHTGSGGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
DB 541 IGAPGEDNRGAVYLPHTGSGGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
QY 601 DLTVGAGHVLRLRSQPLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQK 660
DB 601 DLTVGAGHVLRLRSQPLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQK 660
QY 661 STRDLREGQIOSVVTYDLALDSGRPHSAVFNETKNSTRROTQVGLTQTCETLKLQLP 720
DB 661 STRDLREGQIOSVVTYDLALDSGRPHSAVFNETKNSTRROTQVGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNLSLVGTPLSAFNGLRPLVAEDAQRLLFTALPFPEKNCNGNDNICQDD 780
DB 721 NCIEDPVSPIVLRNLSLVGTPLSAFNGLRPLVAEDAQRLLFTALPFPEKNCNGNDNICQDD 780
QY 781 LSITFSFMSLDCLVGGPREFNTVTVRNDGEDSYRTQVTFPPDLSTYRKVSTLQNR 840
DB 781 LSITFSFMSLDCLVGGPREFNTVTVRNDGEDSYRTQVTFPPDLSTYRKVSTLQNR 840
QY 841 QRSWLACASASTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACASASTEVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA 900
QY 901 NVTSENMPRTNKTPEQLPLVKYAVYVMTVSHGVSTKYLNFNTASNTSRVVMQHOYQVSN 960
DB 901 NVTSENMPRTNKTPEQLPLVKYAVYVMTVSHGVSTKYLNFNTASNTSRVVMQHOYQVSN 960

QY 961 LQORSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPV 1020
DB 961 LQORSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPV 1020
QY 1021 VNCSTAVCORIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTABILFNDVSF 1080
DB 1021 VNCSTAVCORIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTABILFNDVSF 1080
QY 1081 TLLPGQGAFAVSQRTETKVEPEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKROYKD 1140
DB 1081 TLLPGQGAFAVSQRTETKVEPEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKROYKD 1140
QY 1141 MMSEGGPPGAEPQ 1153
DB 1141 MMSEGGPPGAEPQ 1153

RESULT 10

US-08-476-062A-43
Sequence 43, Application US/08476062A
Patent No. 5877275
GENERAL INFORMATION:
APPLICANT: Arnaout, M. Amin
TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
RESPONSES WITH BETA2 INTEGRINS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-476-062A-43

Query Match 99.5%; Score 5925.5; DB 2; Length 1152;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 MALRVLLLTALTLCCHGFNDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVANQR 60

Db 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQSRVVVVGAPQEIIVANQR 60
QY 61 GSYLQCDYSTGSCPTRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVK 120
Db 61 GSYLQCDYSTGSCPTRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVK 120
QY 121 GLCFLFGSNLRQOPQKPFPEALRCCPOEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQL 180
Db 121 GLCFLFGSNLRQOPQKPFPEALRCCPOEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQL 180
QY 181 KSKTTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
Db 181 KSKTTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300
QY 301 NTIASKPPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQBGFSAAIT 360
Db 301 NTIASKPPRDHVQVNNFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQBGFSAAIT 360
QY 361 SNGPLLTSGSYDMAGGVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLV 420
Db 361 SNGPLLTSGSYDMAGGVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLV 420
QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
Db 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480
QY 481 HYIEQTRGGVSVCPPLPRGORARWQCDVLYGQGPWGRFGAALTVLGVDVNGDKLTDVA 540
Db 481 HYIEQTRGGVSVCPPLPRG - RARWQCDVLYGQGPWGRFGAALTVLGVDVNGDKLTDVA 539
QY 541 IGAPGEDNRGAVYLPHTGSGSISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Db 540 IGAPGEDNRGAVYLPHTGSGSISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 599
QY 601 DLTVGAQGHVLLRSQVLRVKAIMEFNPREVARNFECDNVVKGEGEVAVCLHVOK 660
Db 600 DLTVGAQGHVLLRSQVLRVKAIMEFNPREVARNFECDNVVKGEGEVAVCLHVOK 659
QY 661 STDRRLREGIQSVVTVYDLALDSGRPHSRVAFNETKNSRRQTQVGLTTCETLKLQLP 720
Db 660 STDRRLREGIQSVVTVYDLALDSGRPHSRVAFNETKNSRRQTQVGLTTCETLKLQLP 719
QY 721 NCIEDPVSPVLRLNFSLVGTPLSAFCNLRPVLAEADAQRLFTALFPPEKNCNDNICQDD 780
Db 720 NCIEDPVSPVLRLNFSLVGTPLSAFCNLRPVLAEADAQRLFTALFPPEKNCNDNICQDD 779
QY 781 LSITFSPMSLDCLVVGCPREPNVTVYRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNR 840
Db 780 LSITFSPMSLDCLVVGCPREPNVTVYRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNR 839
QY 841 QRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
Db 840 QRSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 899
QY 901 NVTSENNMPRTNKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
Db 900 NVTSENNMPRTNKTEFQLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSN 959
QY 961 LGQSRSLPISLVLPVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
Db 960 LGQSRSLPISLVLPVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1019
QY 1021 VNCISIAVCORIQCDIPFEGIOEEFNATLKGNSLFDWIKTSHNHLIVSTAEILFNDVSF 1080
Db 1020 VNCISIAVCORIQCDIPFEGIOEEFNATLKGNSLFDWIKTSHNHLIVSTAEILFNDVSF 1079
QY 1081 TLLPQOGAFVRSQETETKVEPPEVPNPPLFIVGSSVGGLLLLALITAALYKLGFFKQYKD 1140

Db 1080 TLLPQOGAFVRSQETETKVEPPEVPNPPLFIVGSSVGGLLLLALITAALYKLGFFKQYKD 1139
QY 1141 MMSEGGPPGABPQ 1153
Db 1140 MMSEGGPPGABPQ 1152
RESULT 11
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnsout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US96-01314-43

Query Match 99.5%; Score 5925.5; DB 5; Length 1152;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQSRVVVVGAPQEIIVANQR 60
Db 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQSRVVVVGAPQEIIVANQR 60
QY 61 GSYLQCDYSTGSCPTRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVK 120
Db 61 GSYLQCDYSTGSCPTRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVK 120
QY 121 GLCFLFGSNLRQOPQKPFPEALRCCPOEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQL 180
Db 121 GLCFLFGSNLRQOPQKPFPEALRCCPOEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQL 180
QY 181 KSKTTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
Db 181 KSKTTLFSLMOYSEEPRIHFTFKFQNNPNRSLVKPITQLGRTHATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300

301 NTIASKPPRDHVFQVNNFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
301 NTIASKPPRDHVFQVNNFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
361 SNGPILLSVGSVDWAGGVFLYTSKESKSTFNNTRVDSNDNDAYLGAAAIILNRVQSILV 420
361 SNGPILLSVGSVDWAGGVFLYTSKESKSTFNNTRVDSNDNDAYLGAAAIILNRVQSILV 420
421 LGAPRYQHIGLVAMFRONTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
421 LGAPRYQHIGLVAMFRONTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
481 HYYEOTRGQVSVCPPLRGORARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
481 HYYEOTRGQVSVCPPLRG - RARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 539
541 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
540 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLV 599
601 DLTVGAQGHVLLRSQPLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOK 660
600 DLTVGAQGHVLLRSQPLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOK 659
661 STRDLREGQIQSVVTVYDLALDSGRPHSAVFNETKNSTRRTQVGLTQTCTETLKLQLP 720
660 STRDLREGQIQSVVTVYDLALDSGRPHSAVFNETKNSTRRTQVGLTQTCTETLKLQLP 719
721 NCIEDPVSPIVLRNFSLVGTPLSAFNLRPVLAEDAQRLLFTALPFPEKXKNCNDNICQDD 780
720 NCIEDPVSPIVLRNFSLVGTPLSAFNLRPVLAEDAQRLLFTALPFPEKXKNCNDNICQDD 779
781 LSITFSPMSLDCLVVGGPREFNVTVVRNDGEDSVRTQVTFEPFLDLSYRKVSTIQNORS 840
780 LSITFSPMSLDCLVVGGPREFNVTVVRNDGEDSVRTQVTFEPFLDLSYRKVSTIQNORS 839
841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKA 900
840 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKA 899
901 NVTSENMPRTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 960
900 NVTSENMPRTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 959
961 LQORSUPISLVFLVPLVRLNQTVIWDPRQVTFESNLSTCTKTERLPSSHDFLAEALRKA 1020
960 LQORSUPISLVFLVPLVRLNQTVIWDPRQVTFESNLSTCTKTERLPSSHDFLAEALRKA 1019
1021 VNCIAVCQRIQCDIPFFGIQEBFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSF 1080
1020 VNCIAVCQRIQCDIPFFGIQEBFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSF 1079
1081 TLLPGOGAFVRSQTEKVPFEPVNPPLIIVGSSVGGLLLLALITAALYKLGFKKQYKD 1140
1080 TLLPGOGAFVRSQTEKVPFEPVNPPLIIVGSSVGGLLLLALITAALYKLGFKKQYKD 1139
1141 MMSEGGPPGAEPQ 1153
1140 MMSEGGPPGAEPQ 1152

RESULT 12

5424399-2

; Patent No. 5424399

; APPLICANT: ARNAOUT, M. AMIN

; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS

; NUMBER OF SEQUENCES: 12

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/78,871

; FILING DATE: 16-JUN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 539,842

; FILING DATE: 18-JUN-1990

APPLICATION NUMBER: 212,573
FILING DATE: 28-JUN-1988
SEQ ID NO: 2:
LENGTH: 1152
5424399-2

Query Match
Best Local Similarity 99.5%; Score 5925.5; DB 6; Length 1152;
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARGFQSVVQLQGSRVVVVGAPOEIVAAANOR 60
DB 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARGFQSVVQLQGSRVVVVGAPOEIVAAANOR 60
QY 61 GSLVQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120
DB 61 GSLVQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120
QY 121 GLCFUFGNLRQOPQKFPPEARLARGCPQEDSDIAFLIDGSGSIIPHDFRMRKKEFPVSTMQL 180
DB 121 GLCFUFGNLRQOPQKFPPEARLARGCPQEDSDIAFLIDGSGSIIPHDFRMRKKEFPVSTMQL 180
QY 181 KSKTFLFSLMOYSEBEFRTHFTFKFQNNPNRSLVKPITOLLGRTHATGIRKVVRELFPN 240
DB 181 KSKTFLFSLMOYSEBEFRTHFTFKFQNNPNRSLVKPITOLLGRTHATGIRKVVRELFPN 240
QY 241 ITNGARKNAFKLVITIDGKFGDPLGYEDVIPEADREGVIRYVIGVGDAFSEKSRQEL 300
DB 241 ITNGARKNAFKLVITIDGKFGDPLGYEDVIPEADREGVIRYVIGVGDAFSEKSRQEL 300
QY 301 NTIASKPPRDHVFQVNNFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
DB 301 NTIASKPPRDHVFQVNNFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 360
QY 361 SNGPILLSVGSVDWAGGVFLYTSKESKSTFNNTRVDSNDNDAYLGAAAIILNRVQSILV 420
DB 361 SNGPILLSVGSVDWAGGVFLYTSKESKSTFNNTRVDSNDNDAYLGAAAIILNRVQSILV 420
QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
DB 421 LGAPRYQHIGLVAMFRONTGMWESNANVKTGTOIGAYFGASLCSVDVDSNGSDTLVLIGAP 480
QY 481 HYYEOTRGQVSVCPPLRGORARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYYEOTRGQVSVCPPLRG - RARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVA 539
QY 541 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
DB 540 ICAPGEEDNRGAVYLFHGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLV 599
QY 601 DLTVGAQGHVLLRSQPLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOK 660
DB 600 DLTVGAQGHVLLRSQPLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOK 659
QY 661 STRDLREGQIQSVVTVYDLALDSGRPHSAVFNETKNSTRRTQVGLTQTCTETLKLQLP 720
DB 660 STRDLREGQIQSVVTVYDLALDSGRPHSAVFNETKNSTRRTQVGLTQTCTETLKLQLP 719
QY 721 NCIEDPVSPIVLRNFSLVGTPLSAFNLRPVLAEDAQRLLFTALPFPEKXKNCNDNICQDD 780
DB 720 NCIEDPVSPIVLRNFSLVGTPLSAFNLRPVLAEDAQRLLFTALPFPEKXKNCNDNICQDD 779
QY 781 LSITFSPMSLDCLVVGGPREFNVTVVRNDGEDSVRTQVTFEPFLDLSYRKVSTIQNORS 840
DB 780 LSITFSPMSLDCLVVGGPREFNVTVVRNDGEDSVRTQVTFEPFLDLSYRKVSTIQNORS 839
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKA 900
DB 840 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKA 899
QY 901 NVTSENMPRTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 960
DB 900 NVTSENMPRTNKTBFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQSN 959

QY 961 LGORSLPISLVFLVPVRLNQTIVDRPQVTFSENLSTCHTKERLPSSHDSFLAELRKAPV 1020
DB 960 LQORSPISLVFLVPVRLNQTIVDRPQVTFSENLSTCHTKERLPSSHDSFLAELRKAPV 1019
QY 1021 VNCISIAVCQRIQCDIPFGIOEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSF 1080
DB 1020 VNCISIAVCQRIQCDIPFGIOEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSF 1079
QY 1081 TLLPQCAFVRSQVTFEVEPVPNPLIVGSSVGLLALITAAALYKLGFFKQYKD 1140
DB 1080 TLLPQCAFVRSQVTFEVEPVPNPLIVGSSVGLLALITAAALYKLGFFKQYKD 1139
QY 1141 MMSEGGPPGAEPO 1153
DB 1140 MMSEGGPPGAEPO 1152

RESULT 13

US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-062A-44

Query Match 58.8%; Score 3503; DB 2; Length 1163;
Best Local Similarity 61.3%; Pred. No. 4.2e-287;
Matches 700; Conservative 138; Mismatches 297; Indels 6; Gaps 4;
QY 5 VLLLTALTLCGHNLDENAMTFQENARGFGSVVLOGSRVVGAPQEIIVAAHQRSGLY 64

DB 8 LLLFTALATSLGFNLDTEELTAFTAVDSAGFSDSVVQVYANSWVVGAPQKIITAANQTGGLY 67
QY 65 QCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCTSENYVKGUCF 124
DB 68 QCGYSTGACEPIGLQVPPEAVNMSLGLSLASTTSPQLACGPTVHHECGRNMVLTGLCF 127
QY 125 LFGSNURQOPKFPPEALRCPCPEDSDIAFLIDGSGSIIPHDRRMKEFVSTVMEQLKXSK 184
DB 128 LLGPT--QUTQRLPVSQECPRQEQDIVELIDGSGSISRRNFATMWNFVRAVIFQFORS 185
QY 185 TLPSLMQYSEEFRIHFTKEFQNNPNRSLVXPIITQLGTRHTATGIRKRVARELFNITNG 244
DB 186 TQPSLMQFNKQTHFTFEFRRTSNPLSLLASVHQLQGYTATATAIQNVVHRLFHASYC 245
QY 245 ARKNAPKILVITDGEKFGDPIGYEDVPEADREGVIRVYVIGVDAPFRSEKSRQELNTIA 304
DB 246 ARDATKILVITDGGKKGSDLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSMKELNDIA 305
QY 305 SKPPDRHVQVNNFEALKTIQNLREKIPIAIEGTQGTSSSSPEHMSQEGFSAITNSGP 364
DB 306 SKPSQHIKFVEDFDALQIQNLKEKIPIAIEGTETTSSSSFELEMAQEGFSAVFTPDGP 365
QY 365 LLSTVSGSYDWAGGVFLYTSKEKSTFINMTVDMDMDAYLGYAAAIIILNRNVOSLVLGAP 424
DB 366 VLGAUSFTWAGGAFLLPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVOSLVLGAP 425
QY 425 RYQHIGLVAMFRONTGMESNANVKGITGAYGASLCSVDVDSNGSTDVLVIGAPHYYE 484
DB 426 RYQHTGKAVIFTQVSQRWMAEVTGTQIGSYEGASLCSVDVDTGSDTDLVIGAPHYYE 485
QY 485 QTRGGQVSVCLPRGORARWQCDVLYGEGQWPGFAGALTVLGVDVNGDKLTDVAIGAP 544
DB 486 QTRGGQVSVCLPRGWR--RWCDAVLYGEGQWPGFAGALTVLGVDVNGDKLTDVIGAP 544
QY 545 GEEDNRGAVYLFHGTSGSGISPSHSORISAGSLSPRLQYFGQSLSGQDITMDGLVOLT 604
DB 545 GEENRGAVYLFHGVGLGPSISPSHSQRIAGSLSRLOYFGQALSQGGDITQDGLVDLAV 604
QY 605 GAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVYKGEAGVRVCLHVQKSTRD 664
DB 605 GARGQVLLLRTRPVLVGVSMQFIPABI PRSAFECEQVYVSEQTLVQSNICLYIDKRSKN 664
QY 665 RLREGQISVVTDLALDSDGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIE 724
DB 665 LLASRDQSSVTLDLALDPGLSPRATFQETKRSLSRVRLVGLKAHCENFNLLPSCVE 724
QY 725 DPVSPVLRNLSLVGTPLSAFGNLRPVLAEDAQLFTALPPPKNCNDNICODLSIT 784
DB 725 DSVTPITLRNLNFTLVGKPLLAFLNRLPMLAALQRYFTASLPFEKNCADHICODNLGIS 784
QY 785 FPFMSLDCLVVGVPREFNVTVTVRNDGEDSVYRTQVTFPFLDLSYKRVKSTLQNRQSRW 844
DB 785 FSPFGLKSLVGSNLELNAEVMVWMDGEDSVYTTITFSHPAGLSRYVVAEQKQQLRSL 844
QY 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFMTTFDVSQASLGNKLLKANVTS 904
DB 845 HLTCDSPVGG--SQGTWSTSCRINHLPFRGQAQITFLATFDVSPKAVLDGRLLTANVSS 902
QY 905 ENNMPTNKTEFOLELPVKYAVVMVTSVSHGVSTVYLNFTAS--ENTSRVMOHOYOVSNLQ 963
DB 903 ENNTPRTSKTTFQLELPVKYAVYTVVSSHEQFTKLPFSESEKSHVAMRYQVNNLQ 962
QY 964 RSLPISLVFLVPVRLNQTIVDRPQVTFSENLSTCHTKERLPSSHDSFLAELRKAPVVC 1023
DB 963 RDLPVSNFVWYVVELNQEAVMMDVEVSHQPNSLRCSSEKAPPASDPLAHIQKNPVLDC 1022
QY 1024 SIACVQRIQCDIPFGIOEEFNATLKNLSFDWYIKTSHNHLIVSTAEILPNDVSFTLL 1083
DB 1023 SIAGCLRPDCVPSFVQOEELDTLKNLSFGWVRQILQKKVSVVVAEITFDTSVSQL 1082
QY 1084 PQCAFVRSQVTFEVEPVPNPLIVGSSVGLLALITAAALYKLGFFKQYKDNMS 1143
DB 1083 PQCAFVRSQVTFEVEPVPNPLIVGSSVGLLALITAAALYKLGFFKQYKDNMS 1142

Qy 1144 E 1144
Db 1143 E 1143

RESULT 14

PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-44

Query Match 58.8%; Score 3503; DB 5; Length 1163;
Best Local Similarity 61.3%; Pred. No. 4.2e-287;
Matches 700; Conservative 138; Mismatches 297; Indels 6; Gaps 4;
Qy 5 VLLLTALTLCHGNLDTEENAMTQENARGQSVVQLQSGSRVVGAPQEIIVANQSGSLY 64
Db 8 LLLLTALTSGLGNLDTEELTAFRVDSAGGDSVVQYANSVVVGAPQKITAAANQGGLY 67
Qy 65 QCDYSTGCEPIRLQVPEAVNMSLGLSLAATSPPOLACGPTVHGTCSNTYVXGLCF 124
Db 68 QCGYSTGACPIGLQVPEAVNMSLGLSLASTTSPSOLLACGPTVHCEGRNMYLTGLCF 127
Qy 125 LFGSNLRQQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKS 184
Db 128 LLGPT--QLTQRLPVSRQECPRQEQDIVFLIDGSGSISSRNFMATMNFVRAVISQFORPS 185
Qy 185 TFLSLMOYSEPIRHFTFKFQNNPNRSLVKIPIITOLLGTHATGIRKVVRELFNITNG 244
Db 186 TQSLMQFSNKFQTHFTFEERFSTNPLSLASVHQLQGFYATATQVNVHRLFHASYG 245
Qy 245 ARKNAFILVITDGEKFGDPLGVEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTIA 304
Db 246 ARDATKILVITDGEKFGDPLGVEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTIA 305
Qy 305 SKPPRDRHVQNNFEALKTQNLREKIFAIEGTQTGSSSSSFEHMSQEGFSAAITNGP 364

Db 306 SKPSQEHIPKVEDFDALKDIOQLKEKIFAIEGTQETTTSSSSFELEMAQGFSAVFPDGP 365
Qy 365 LLSTVGSYDWAGGVFLYTSKEKSTFINMTVRVSDMDNDAYLGAAAAIILNRVQSLVGLAP 424
Db 366 VLGAVGSGFTWGGAFLYPPNMSPTFINMSQENVMDSDSYLGSTELALWKGVQSLVGLAP 425
Qy 425 RYQHIGLVAMFRONTQMWSNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYYE 484
Db 426 RYQHTGKAVIFTQVSRQWRMKAETGTQIGSYFGASLCSVDVDTGSDTLVIGAPHYYE 485
Qy 485 QTRGGGVSVCPPLPRGRABWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAP 544
Db 486 QTRGGGVSVCPPLPRGWR--RWWCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVVGAP 544
Qy 545 GBEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTV 604
Db 545 GREENRGAVYLFHGVLGPSISPSHSQRIAGSQSLSRLOYFGQALSGQDLTQDGLVDLAV 604
Qy 605 GAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDNVQVKGKEAGEVRLVHVQKSTRD 664
Db 605 GARGQVLLLRTRPVLWVGVMQFIPAEIPRSAFECEQVQVSEQTLVQSNICLYIDKRSKN 664
Qy 665 RLREGQIOSVTVYDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIE 724
Db 665 LLGSRDLQSSVTLDLALDPGRLSPRATFOETKNSRSLRVRLGLKAKHCENFLLPSCVE 724
Qy 725 DPVSPITVRLNFSLVGTPLSAFQNLRPVLAEDAQRLLFTALFPFEKNCNDNICQDLSIT 784
Db 725 DSVTPITLRLNFTLVGKPLAFRNLRLPMLAALAAQRYFTASLFEKNCNGADHICQDNLGIS 784
Qy 785 FSPMSLDCLVVGCPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWS 844
Db 785 FSPFGLKSLLVGSLNLENAEVMVMNDGEDSYGTTITFSPAGLSYRYVAEGOKQGLRSL 844
Qy 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTS 904
Db 845 HLTCDSAPVG--SQGTWSTSCRNHLIFRGGQITFLATFDVSPKAVLGDRLLLTANVS 902
Qy 905 ENNMPTNKTEFQLELPVKYAVVMVVTSHGVSTKYNLFTAS--ENTSRVMOHYQVSNLQ 963
Db 903 ENNTPTSKTTFQLELPVKYAVTVVSSHEQTKYLNFSSEKESHVAMHRVQVNNLQ 962
Qy 964 RSLPISLVLPVRLNQTVIMDRPQVTFSENLSSTCHTKERLPFSHSDFLAELRKAPVNC 1023
Db 963 RDLPSVSNFVWVPELNQEAVMDEVSHPQNPRLCSSEKSIAPPASDFLAHIQKNPVLDC 1022
Qy 1024 STAVCORIOCDIPFFGIEEFNATLKGNSLSPDWYIKTSHNHLIIVSTAILFNDSVFTLL 1083
Db 1023 SIAGCLRFCDVPFSVQBEELDFTLKGNLSFGVVRQILQKQKSVSVSVAEITDTSVSQL 1082
Qy 1084 PGQGAFFVRQOTETKVPPFVPPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQDMS 1143
Db 1083 PGQEAFFMRAQTTVLEKYKVNPTPLIVGSSIGGLLLALITAVLYKVGFFKQYKQEMME 1142
Qy 1144 E 1144
Db 1143 E 1143

RESULT 15

US-08-173-497-4
; Sequence 4, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60608-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,497
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 543795band, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31363
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-173-497-4

Query Match 58.4% Score 3480; DB 1; Length 1163;
Best Local Similarity 61.1% Pred. No. 3.7e-285;
Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;
QY 5 VLLTALTLCHGFLNLTENAMTFQENARGQSVVLOGSRVVGAPQEIIVAAQSGSLY 64
DB 8 LLLFALATSLGFLNLTDELTAFRVDSAGFGDSVQVYANSWVVGAPQKIIAANQIGGLY 67
QY 65 QCDYSTGCEPIRLQVPEAVNMISGLSLAATTSPQLACGPTVHOTCSENTVYKGLCF 124
DB 68 QCGYSTGACEPIGLQVPEAVNMISGLSLAATTSPQLACGPTVHCEGGRNWTGLCF 127
QY 125 LFGSNLRQOPKQFPALRGCPQEDSDIAFLIDGSGSIIPHDPFRMKFVSTVMEQLKSK 184
DB 128 LLGPT--QLTQLPVSQCEPQEQDIIVFLIDGSGSISSRNFAFMNFRAVISQFQPS 185
QY 185 TLFSLMQYSEFRIHFTKPEQNNPNRSLVKPITQLGRTHATGIRKVVRELPNITNG 244
DB 186 TQFSLMQFSNKFQTHFTFEFRRTSNPLSLASVHQLOQFTVYATAIQNVVHRLPHASYG 245
QY 245 ARKNAFKILVITDCEKGDPLGEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIA 304
DB 246 ARRDIAKILVITDCKGSDSLDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKLNDIA 305
QY 305 SKPPRDHYFQVNNFALKTIQNLREKIFAIBGTQTSSSSPEHEMSQEGFSAATISNGP 364
DB 306 SKPSQEHFKVEDFDALQIQNLKEIFAIBGTETISSSSFELEMAQEGFSAVTPDGP 365
QY 365 LLSTVGSYDWAAGVFLYTSKSTFTINMTRVDSNDNDAYLGYAAAIILRNRVQSLVGLAP 424
DB 366 VLGAAGVFTWSGCAFLYPPNMPPTINNSQENVMDRDSYLGYSTELALWKGVQSLVGLAP 425
QY 425 RYQHIGLVAMPQNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYE 484
DB 426 RYQHIGKAVIFTQVSRQRMKAETVGTQIGSYFGASLCSVDVDTDGSSTDVLIGAPHYYE 485
QY 485 QTRGGQSVCPILPRGORARWQCDVLYGEOGQWGRFGAALTVLGDVNGDKLTDVAIGAP 544
DB 486 QTRGGQSVCPILPRGWR--RWCDVLYGEOGQWGRFGAALTVLGDVNGDKLTDVIGAP 544
QY 545 GEEDNRGAVYLFPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTV 604
DB 545 GEENRGAVYLFHGVLPSPISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTQDGLVDLAV 604

QY 605 GAQGHVLLRSQPLRVKAIMFENPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRD 664
DB 605 GARGQVLLLRTRPVLWVGYSMQFIPAEIPRSAFECEQVYVSEQTLVQSNICLYIDKRSKN 664
QY 665 RLREGIOISVYTDLALDSGRPHSRVAFNETKNSRQTOVGLTGTOTCETLKLQOLNCIE 724
DB 665 LLGSRDLQSSVTLDLALAPGRLSRAIFOETKRSLSRVRLVGLKAHCENFNLLPSCVE 724
QY 725 DPVSPITVRLNLSLVTPLSAFGLNRPVLAEDAQRFLTALFFPEKCKGNDNICODLSIT 784
DB 725 DSVIPILRLNFTLVGKPLLAFLNRLPMLAALQRYFTASLPEKCKGADHICODNLGIS 784
QY 785 FFSMSLDCLVVGQPRBFNVTVVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQORSW 844
DB 785 FSPFGLKSLVGSNLELNAEVMVMWMDGDSYGTITFSPAGLSYRVVAEGQKQGLRSL 844
QY 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTS 904
DB 845 HLTC--CSAPVGSQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVGLDRLLLIANVSS 902
QY 905 ENNMPTNKTEFOLELPVKYAVYVMVTSHGVSTKYLNFAS--ENTSRVMOHQYQVSNLQ 963
DB 903 ENNIPRTSKTIFOLELPVKYAVYVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQ 962
QY 964 RSLPISLVFLVPLVRLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNC 1023
DB 963 RDLPVSIWFWVVELNQEAVMMDVEVSHQPQNSLRCSSEKIAPPASDLAHIQNPVLDC 1022
QY 1024 SIACQRIQCDIPFFGLOEEFNATLKNLSFDWIKTSHNHLILVSTAEILFENDSVFTLL 1083
DB 1023 SIAGCLFRCDVPSFSVQEEELDTLKNLSFGWVROQLQKKVSVVSAEIIFDTSVYSQ 1082
QY 1084 PGQAFVRSOTETKVEPFPVNPPLPLIVGSSVGLLALLALITAAALKYKLGFPKRYKDMMS 1143
DB 1083 PGQAFVRSOTETKVEPFPVNPPLPLIVGSSVGLLALLALITAAALKYKLGFPKRYKDMMS 1142
QY 1144 E 1144
DB 1143 E 1143

Search completed: November 25, 2003, 14:23:09
Job time : 17.7527 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 14.9579 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 FNLDTENAMTFQENARGFQ.....FKRQYKDMSEGGPFGEAPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying Chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum-Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	1 RWHU1B	cell surface glyco
2	4447	75.6	1153	2 S00551	leukocyte surface
3	3456	58.8	1163	1 RWHU1C	cell surface glyco
4	1532.5	26.1	1170	2 S03308	cell surface glyco
5	1516.5	25.8	1163	2 I56126	lymphocyte fuction
6	1128	19.2	1179	2 A53213	integrin alpha-E c
7	1085.5	18.5	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45914	integrin alpha 2 s
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1054	17.9	1181	2 A33998	integrin alpha-2 c
11	1049	17.8	1180	2 A35854	integrin alpha-1 c
12	666	11.3	1039	2 A41131	lymphocyte-Peyer's
13	642	10.9	1038	2 S06046	integrin alpha-9 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha cha
16	579.5	9.9	1054	2 JCT294	alphan integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.4	1053	2 S44250	integrin alpha-5 c
20	551.5	9.4	1034	2 A36108	integrin alpha-V c
21	539	9.2	1044	2 T10050	integrin alpha-v c
22	533.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	fibronectin recept
24	531	9.0	1073	2 B36429	integrin alpha-6 c
25	530	9.0	1048	2 A27421	integrin alpha-5 c
26	529.5	9.0	1051	2 A40021	integrin alpha-3 c
27	524.5	8.9	1091	2 A41543	integrin alpha-6 c
28	514.5	8.8	1044	2 S16516	integrin alpha-8 c
29	497	8.5	1394	2 A29637	position-specific

RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: Complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A45526; A26091; I52567

J:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

R:Biochem. J. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:Cross-references: GB:J03925; NID:gl87284; PIDN:AAA59544.1; PID:g307148

A:Note: Part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J:Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499, 501-965, 'P', 967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAA5

A:Note: The authors translated the codon TAC for residue 1129 as Thr

A:Note: Part of this sequence, including the amino end of the mature protein, was con

R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:gl80018; PIDN:AAA58410.1; PID:g553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe

A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:Cross-references: GB:J04145; NID:gl89068; PIDN:AAA59903.1; PID:g386975

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526; MUID:93123748; PMID:8419480
A:Accession: A46526
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499,501-1153 <FL>
A:Cross-references: GB:52227; NID:9263047; PIDN:AAB24821.1; PID:g263049
A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
A>Note: sequence extracted from NCBI backbone (NCBIP:121963)
R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26091
A:Molecule type: protein
A:Residues: 17-31 <PI>
A:Experimental source: granulocytes
R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A:Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: 152567; MUID:92144986; PMID:1346576
A:Accession: 152567
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:Cross-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219
C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genet.Cs:
A:Gene: GDB:ITGAM; CR3A
A:Cross-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; ma
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F:17-1108/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <WAA>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.3% Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9% Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFQSVVQLQSGRVVVGAPQEIIVANQKSLYQCDYSGSCEPI 60
Db 17 FNLDTENAMTQENARGFQSVVQLQSGRVVVGAPQEIIVANQKSLYQCDYSGSCEPI 76

Qy 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFPGSNLRQOPQK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFPGSNLRQOPQK 136

Qy 121 FPEALRGCPEDSDIAFLIDSGSIIPHDFRMKELVSTIMEOLKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPEDSDIAFLIDSGSIIPHDFRMKELVSTIMEOLKSKTLFSLMOYSEEF 196

Qy 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVVLGFGDAFRSEKSKQLNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRYVVLGFGDAFRSEKSKQLNTVASKPPRDHVFQVN 316

301 NFEALKTQVQLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLSTVGSYDWAG 360
Db 317 NFEALKTQVQLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLSTVGSYDWAG 376

Qy 361 GVFLYTSKSKSPFINNTRVDSMDNDAYLGYAAAILRNVRQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKSKSPFINNTRVDSMDNDAYLGYAAAILRNVRQSLVLGAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSCPL 480
Db 437 QNTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSCPL 496

Qy 481 PRQQRARWQCDVLYGCEQGPWGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAYILF 540
Db 497 PRQQRARWQCDVLYGCEQGPWGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAYILF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

Qy 601 PVLVRKAIIMEFNPREVARNVFECDQVVGKEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFECDQVVGKEAGEVRVCLHVOKSTRDLRREGQIOSVVT 676

Qy 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 736

Qy 721 SLVGTPLSAFNGNLRPVLAEDAQRLFTALPFFKNCNGNDNICQDDLSITFSFMSLDCLVG 780
Db 737 SLVGTPLSAFNGNLRPVLAEDAQRLFTALPFFKNCNGNDNICQDDLSITFSFMSLDCLVG 796

Qy 781 GPREFNVTVVRNDGSDSYRTQVTFEFPDLDSYRKYSTLQNSQRSWRACASASSTEV 840
Db 797 GPREFNVTVVRNDGSDSYRTQVTFEFPDLDSYRKYSTLQNSQRSWRACASASSTEV 856

Qy 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916

Qy 901 QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVQHQYQVNSLGCORSIPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVQHQYQVNSLGCORSIPISLVFLVPV 976

Qy 961 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISAVCORIQCDIP 1020
Db 977 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISAVCORIQCDIP 1036

Qy 1021 FFGTQBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGFVRSQTET 1080
Db 1037 FFGTQBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGFVRSQTET 1096

Qy 1081 KVBPFEPVNPPLPIVSGVGLLLALITAAALYKLGFFKRYKQKDMSEGPPGABEQ 1137
Db 1097 KVBPFEPVNPPLPIVSGVGLLLALITAAALYKLGFFKRYKQKDMSEGPPGABEQ 1153

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N:Alternate names: complement-3 receptor alpha chain
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
C:Accession: S00551; 159078
R:Pytela, R.
EMBO J. 7, 1371-1378, 1988
A:Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with th
A:Reference number: S00551; MUID:98312584; PMID:3044779
A:Accession: S00551
A:Molecule type: DNA
A:Residues: 1-1153 <PYT>
A:Cross-references: EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52983
A>Note: the authors translated the codon CAC for residue 569 as Gln

R;Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
A:Reference number: I59078; MUID:86287312; PMID:2942940
A:Accession: I59078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 75.6%; Score 4447; DB 2; Length 1153;
Best Local Similarity 73.5%; Pred. No. 2.1e-302;
Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARFGQSVVLOGSRVVVCAPOEIVAAANORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTEHPMTFQENAKFGQNVVQLGCTSVVAAPOEAKAVNQTGALYQCDYSTSRCHPI 76

QY 61 RLQVPVAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSLNRQQOK 120
DB 77 PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQCKENTYVNGCYLFGSLNLRPPQO 136

QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHDFRMRKELVSTIMEQKKSKTLFSLMQYSBEF 180
DB 137 FPEALRECPQESDI VFLDGGSGINNIDFQMKKEFVSTVMEQKKSKTLFSLMQYSDEF 196

QY 181 RHFTFKFQNNPNPSLKPITQLLGRTHATGLRKVRRELPNITNGARKNAFKILFL 240
DB 197 RHFTFNDKRNPSRSHSPIQLNGRTKTASGIRKVRRELPNITNGARKNAFKILVVI 256

QY 241 TDCEKFGDPLGYEDVPELDREGVIRYVLGFGDAPFSEKSRQELNVTASKPRDHDVFOAN 300
DB 257 TDCEKFGDPLGYEDVPELDREGVIRYVLGFGDAPFSEKSRQELNVTASKPRDHDVFOAN 316

QY 301 NFPAKTVQNLREKIFAIEGTQTSSSFEHEMSQEGFSAAITSNGPLLSVTSVGSVDWAG 360
DB 317 NFPAKTVQNLREKIFAIEGTQTSSSFEHEMSQEGFSAAITSNGPLLSVTSVGSVDWAG 376

QY 361 GVFLYTSKEKSTINTRVDSMDNDAYLVAAAILRNVRVQSLVGLGAPRYQHIGLVAMPR 420
DB 377 GAFLYTSKOKVTFINTTRVDSMDNDAYLVAAAILRNVRVQSLVGLGAPRYQHIGLVAMPR 436

QY 421 QNTGMVESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQSVCP 480
DB 437 ENFGTWEPTSHIKGSIQSGYFGASLCSVDMDADGNTNLILIGAPHYEEOTRGQSVCP 496

QY 481 PRQARWQCDALVYGEQOPWRCFCAALTLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRG-RARWQCEALLHGDQHPWRCFCAALTLDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

QY 541 HGTSGSGISPHSHQRTAGSKLSPRLQYFGQSLGGQDLTDWGLVDLTGVAQHVLRLRQ 600
DB 556 YGASIASLSASHSHRIIGHFSPGLQYFGQSLGGQDLTDWGLVDLTGVAQHVLRLRQ 615

QY 601 PVLRVKAIMEFNPREVARNVECDVWKGAGEVRVCLHVOKSTFDRLRREGQIOSVVT 660
DB 616 PVLRVKAIMEFNPREVARNVECDVWKGAGEVRVCLHVOKSTFDRLRREGQIOSVVT 675

QY 661 YDLALDSGRPHSAFVNETKNSRRTQTVLGLTQTCETLKLQPNCEIPVSPVILRLNF 720
DB 676 YDLALDPVRSRIRAFDEFKNTRTQVFLMQKCEETLKLILPDCVDVDSVPILRLNY 735

QY 721 SLVGTPLSAFGLNRPVLADAEORLFTALFPFKEKNCNDNICQDDLSITFFMSLDCLVVG 780
DB 736 TLVGEPLRSGNLRPVLADAEORLFTALFPFKEKNCNDNICQDDLSITFFMSLDCLVVG 795

RESULT 3

RWUIC

cell surface glycoprotein CD11c precursor - human

N:Alternate names: leukocyte adhesion receptor p150.95 alpha chain

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999

C:Accession: A36584; A35543; S00864

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 12750-12751, 1990

A:Reference number: A36584

A:Contents: erratum

A:Accession: A36584

A:Molecule type: DNA

A:Residues: 1-1163 <COR>

A:Note: this revision to the sequence from reference A35543 includes the carboxyl end

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150.95 molecule

A:Reference number: A35543; MUID:90153306; PMID:2303426

A:Accession: A35543

A:Molecule type: DNA

A:Residues: 1-834 <CO2>

A:Note: this sequence has been revised in reference A36584

R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.

EMBO J. 6, 4023-4028, 1987

A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte

A:Reference number: S00864; MUID:88166645; PMID:3327687

A:Accession: S00864

A:Molecule type: mRNA

A:Residues: 1-755, 'L', 757-1163 <CO3>

A:Cross-references: GB:M81695; EMBL:Y00093; NID:9487829; PIDN:AAA59180.1; PID:g4687830

A:Note: part of this sequence was confirmed by protein sequencing

C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on

C:Genetics:

A:Gene: GDB:ITGAX; CD11C

A:Cross-references: GDB:119758; OMIM:151510

A:Map position: 16p11.2-16p11.2

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology

C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>

F:20-1107/Domain: extracellular #status predicted <EXT>

F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>

F:1108-1133/Domain: transmembrane #status predicted <TM>

F:1134-1163/Domain: intracellular #status predicted <INT>

F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status F

Qy	175	QYSEEPRIHPTKEFQNNP	RSILKPI	TOLLGR	TH	TATGLR	KVRELF	NITNGAR	KNAF	234
Db	197	QFSTSYKTEBDFSDYKWK	KDPA	LLKHXK	MLLL	TNTFGA	INTVATE	VEFREELGAR	PDAT	256
Qy	235	KILFLLTDGKFGDPLGYE	VI	PELDREG	VR	VLGFGD	AFREK	SGROEL	NITVASKPPRD	294
Db	257	KVLIIITDGE--ATDS	NI	DAKD----	II	RIIG	IKH	FQTKESQET	LHKFAS	309
Qy	295	HVFOANNFEALKT	VQNLRE	KIP	FAIB	ETQ	TGSSSE	HEHMSQ	BFGSAAIT	354
Db	310	FKVILDTFEXLK	KDLFTE	LOK	YIV	EGTS	QD	LTSFNM	ELSSSGISADLS	369
Qy	355	SYDWAGGVF-LY	TSKES	TFIN	TRVDS	DMND	AYL	GYAAA-I	ILNR	412
Db	370	AKDWAGGFLD	KADLQD	DT	IGNE	PLTP	VRAG	LYTV	TWLSR	429
Qy	413	IGLVAMFR--ONT	GMES	NSAN	VKGT	IGAY	FGAS	LCSD	VDSNG	470
Db	430	MGRVLLFOEP	QGGH	W	SOVQ	TI	HGT	IGSY	FGELC	489
Qy	471	RGGOVSVCP	PRGOR	AR	WQ	DAV--	LYGEO	QWGR	GAAL	528
Db	490	RGRVFIY----	OR	LOG	EE	VE	SELQ	GD	PGYPL	544
Qy	529	GEEDNRGAV	L	FHGT	SG	SGIS	P	SHSQ	RIAGS	588
Db	545	LEE--QCAV	I	ENGR	HG-	GL	SP	QS	RIEG	601
Qy	589	GAQGHVLL	LSR	Q	P	VL	R	KA	T	647
Db	602	GAESQMT	VS	SR	P	VD	MT	LS	MF	660
Qy	648	DLRLREG	IO	QSV	VT	Y	LD	AL	D	707
Db	661	POP-QG	FL	AN	LY	T	Y	LD	G	718
Qy	708	EDVP	S	P	I	V	L	R	N	757
Db	719	QDLIS	P	N	V	S				

RESULT 5
I56126
lymphocy

C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1
C:Accession: I56126
R:Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A:Title: Cloning of the murine lymphocyte function-associated molecule-1
A:Reference number: I56126; MUID:91268576; PMID:2051027
A:Accession: I56126
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1163 <RES>
A:Cross-references: GB:M60778; NID:g198785; PIDN:AAA39426.1; PID:g198786
C:Genetics:
A:Gene: LFA-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type IV
F:151-315/Domain: von Willebrand factor type A repeat homology <VWA1>

I56126 lymphocyte fuction-associated molecule-1-alpha - mouse

A;Accession: A45226

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1151 <BRI>

A;Experimental source: hepatoblastoma cell line HepG2

A;Note: sequence extracted from NCBI backbone (NCBIP:124326)

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F;142-317/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 18.5%; Score 1085.5; DB 2; Length 1151;
Best Local Similarity 27.6%; Pred. No. 3.1e-67;
Matches 342; Conservative 209; Mismatches 488; Indels 201; Gaps 44;

QY 1 FNLDTENAMTFQENARG-FOQSVVOL---QGSRVVVGAPQEIIVAAQNRGLSYQCDYSTGS 56
DB 1 FNVDKNSMTFGPVEDMFCVTVQYENEBGKVLIGSLVGPQKRTGDVYKCPVGRGE 60

QY 57 CEP-IRLOVPEA-----VMSLGLSLAATTSPQQLACGPTVHTQCSNTYVKG 106
DB 61 SLPCVKLDLPVNTSIPNVTVEKNMTFGSTL-VTNPNGGFLACGPLYAVRCGHLHYTTGI 119

QY 107 CFLFGSNLRQPOKFPFALRGCPQEDSDIAFLIDGSGIIPHPFRMKELVSTIMBOLKK 166
DB 120 CSDVSPTFQVNSIAP--VOECSTQ-LDIVIVLDGNSIYPWD-----SVTAFELNLLKR 171

QY 167 -----SKTFLSLMOYSEFRIHFTKFEQNNPNRSLIKPITQLGR-THTATGLRKVR 220
DB 172 MDIGPKQTQVGIQVQGNVTHEFNLNKYSSTEEVLVAAKKIVQGRQMTALGTDARTK 231

QY 221 ELFNITNGARKNAFKILFLTDGKFGDPLGYEDVPELDRGVIRYVLGFGDAFR---- 276
DB 232 EAFTARGARGVKVNVIVTDGESH--DNHRLKKVQDCDENIQRFSAIILGSYRGNL 290

QY 277 -SEKSRDELNTVASKPRDHVFQANFEALKTQVONREKIFAIEGTQTGSSSFEHMS 335
DB 291 STEKFEIKSIASEPTKEHFNVSDELAVTIVKTIGERIFALEATADGSAASFEMMS 350

QY 336 QEGFSAITSNGPLLSVSGYDMAGGVFLYTSKE-----KSTF-INNTRVDSMDNAYLG 389
DB 351 QTGFSAHYSQDWMLGAVGAYDNNGTVVMQKASQIIPRTTENVSTKKNEL-ASYLG 409

QY 390 YAAAIILNRVQSL-VLGAPRYOHIGLVAMFRONTGMWESNANVKTQIGAYGASLCV 448
DB 410 YTVNSATASSGDVLYTAGPRYNTGQVVIYRMEDGNKIKILQLSGEIGSYFGSLTUT 469

QY 449 DVDSNGSTDLVLGAPHY-----YEOTR-OGQVSVCLPRGQARQWQDA 492
DB 470 DIDKSDTILLVGAPMYMGTKEEKGKVVVYALNQTFRFYQMSLEPIKOTCCSSRQHNS 529

QY 493 VLYGEOQOPWG-RFGAALTIVLGDVNGDKLTDVAIGAPGEDNDRGAVYLFHGTSGSGISPS 551
DB 530 CTTENKNEPCGARFGTAIAVKDLNLDGFNDIVIGAPLEDHGGAVYIYG-SGKTIRKE 588

QY 552 HSQRIAGSKLSPRLQYFGSLSGQDITMDGLVDLTVGAQGHVLLRLRSQPLVAKAMEF 611
DB 589 YAQRIPSGGDKTLKFPFGSIHGEMDLNGDGLTDVTIGLGGALFWSRDVAVVKTWNF 648

QY 612 NPREVARNVFECNDQVVKGEAG--EVRVCLHQV-KSTRDLREGQIQSVVTDALDSG 668
DB 649 EPNKVNITQKNCH---MEGETVCINATVCFEVKLKSKEDTIYEADLQ----YRVLDL 701

QY 669 RPHSRVAFNET-----KNSTRQTOVLGLTQTCETLKLQLPNCIEDPVPSPVLRNLSL 722
DB 702 RQISRSFSGTQERKVOBNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFNL 753

QY 723 VGTPLSAGNLRPLVLAEDAQRLETFALPFPKNGCNDNICQDDLSITFSFMSLCLVVGGP 782
DB 754 T-DPENG-----PVLDDSLNSVHEYIPFAKDCGKCEKICISLSLHVATTEKDLIIVR 807

QY 783 RE-FNVITVTRNDEGDSRYQTQVFFPDLDSYRKVSTLQNRQSRWRPLACESASSTEV 841
DB 808 NDKFNVSLTVKNTKDSAYNRTTIVHYGPNLVFSGIEAIQKD-----SCESN----- 853

RESULT 8

I45914

Integrin alpha 2 subunit - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999

C;Accession: I45914

J;Kamata, T.; Puzon, W.; Takada, Y.

J;Biol. Chem. 269, 9659-9663, 1994

A;Title: Identification of putative ligand binding sites within the I-domain of integr

A;Reference number: A54402; PMID:94193647; PMID:7511592

A;Accession: I45914

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1170 <KAM>

A;Cross-references: GB:L25886; NID:9439695; PIDN:AAB59255.1; PID:9439696

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F;161-336/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.2%; Score 1071; DB 2; Length 1170;

Best Local Similarity 27.5%; Pred. No. 3.3e-66;

Matches 334; Conservative 218; Mismatches 495; Indels 168; Gaps 48;

QY 1 FNLDTENAMTFQ-ENARGFQSVVOL---QGSRVVVGAPQEIIVAAQNRGLSYQCD--DYST 54

DB 19 YNVGLPKAKIFSGPSSEQFQYAVQVQFINPKGNWLLVGSWSPGPKNRMGDVYKCPVDLST 78

QY 55 GSCEPIRLO-----VPVEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVKG 107

DB 79 TTCEKLNLOTSTSMNSVNTKMTNMSLGLTLTRNVGTGGFLTCGPLWAQCGSQYTTGVC 138

QY 108 FLFGSNLRQPOKFPFALRGCPQEDSDIAFLIDGSGSIIPHPFRMKELVSTIMEOLK-- 165

DB 139 SDVSPDF-QLRISPAVQTCF-SFIDVVVCDENSIYPWD--AVNFKLFVQGLDIG 194

QY 166 KSKTLESLMOYSEEFRIHFTKFEQNNPNRSLIKPITQLL-----GRTHATGLRKVR 221

DB 195 PTKTMGLTIQYANNPRVFNLTNFTFSKD---EMIKATSTQTFQYGGDLTNTFKAIQVARDT 251

QY 222 LFNITNGARKNAFKILFLTDGKFGDPLGYEDVPELDRGVIRY---VLGF--GDAFR 276

DB 252 AYSTAAGRGPGATKVMVVTVDGESH--DGSKLKAVIDQCKNDLTFRGIAVLGLVLYNRNALD 310

QY 277 SEKSRDELNTVASKPRDHVFQANFEALKTQVONREKIFAIEGTQTGSSSFEHMSQ 336

DB 311 TKNLKIEIKAIASIPTEHFHFNVSDEADLLEKAGTIGEQIFSGIEGTQVQ--GDNFQHEMSQ 369

QY 337 EGFSAAIT--SNGPLLSTVGSVDWAGGVFLYTSKEKSTFINMT--RVDSDMN--DAYLGYA 391

Db 370 VGFAEYSPONNLMGLGAVGAYDWSGTVVQKTPHGLHIFSKQAPEQILQDRNHSYLGYS 429
Qy 392 AAILLRNRVQSVLGCAPRYQHIGLVAMFRONTGWNESNANV-----KGTQIGAYFCASL 445
Db 430 VASISTGNSVHFVAGAPANTYTGQIVLYSVN-----ENGNTVVIQSQRGDDIGSYFQSVL 484
Qy 446 CSVDVDSNGSDLVILGAPHYVEQTR--GGQVSVCPPLPRGQARWQCDVAVLYGQGGQPWG 503
Db 485 CAVDVNRKDTITDVLVVGAPMYMNDLKKEGRVYLFTITKG--ILNWH--QFLEGPNGLENA 541
Qy 504 RFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVYLFHGTSGSGISPSHSORIGAS--KL 561
Db 542 RFGSAIALSDINMDGDFNDVVGSPLENQNSGAVIYNGHEGM--IRLRYSKILGSDRAF 600
Qy 562 SPRLOVFCQSLSGGDLTMDGLVDLTGCAQGHVLLRQSVLRYKAIWEPNPREVARNVP 621
Db 601 SSHLOVFRSLDGVGDLNGDSITDVSAGFAQVQVQLWSQSADVSQASFTPKKI--TLL 658
Qy 622 ECNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVYDIALD----SGRPHSRVAFN 677
Db 659 NKNAEI-----KLKLCF----SAKFRPTNQNNQVAIVYNTIDEDQFSSRSVIRGLFK 707
Qy 678 ETQNSTRTQVGLGTQCE--TLKLQLPNCIEDPVSIVLRNPSL--VGTPLSAGNL 733
Db 708 ENNERCLOKTMVSAQRCSEYIIHQEPS---DIISPLNLCMNISLENPGT----- 756
Qy 734 RPLVAEDAQRFLTALFPFKCKGNDNICQDDLSITF-----SFMSLDCLVGGPREFNVTV 789
Db 757 NPALBAYSEVTKVFSIPFKDGGDDGVCISDLVNVQQLPATQQOPFIVSNQNRLLTFSV 816
Qy 790 TVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASST--EVSGALKSTS 848
Db 817 QLKNNKESAYNTEIIVDFSENLPF-----ASMSMPVDGTEVTCQIASSQSVT 864
Qy 849 CSINHIPIPEENSEVTNFTFDVDSKASLGKLLKANVTSENMPRTNKTETQLELPVKY 908
Db 865 CNVGYPALKSQKQVTFTEFNDFNLQ--NLQNAISISFRALSESQEBNADNVSNLKSLLY 923
Qy 909 AVYVMTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQOR-----SLPISLFLV 958
Db 924 DAETHIT--RSTNINFEVSLDGNVSSV--HSFE--DIGPKFISIKVTTGVSVPVMSA--- 976
Qy 959 PVRLNQTVIWRPQVTFSEN--LSSTCHTKE-----RLPSHSDFLAB- 998
Db 977 -----SVLIHPQYTKDNPLMYLTGVHTDQAGDISCEAEINPLKIQTSSTSSVSFKSEN 1030
Qy 999 LRKAPVNCSTAVCQRIQCDIPFGIOBEFNATLKNLSFDWYIKTSHNHLIVSTABI- 1057
Db 1031 FRHIKELNCRATASCNIMCWLRLQVKGVEFLNVSTRIWNGTFAASTFTQVLTAAABID 1090
Qy 1058 LFNDSVFTL-----LRGQAFVRSQETKVEPPE--VPNPLPLIVGSSVGGLLALLITA 1110
Db 1091 TYNPOIYVIENTVVTIP-----LTIMPKHEKVEVPTGVIVGSGVIAGILLALLA 1140
Qy 1111 ALYKLGFFKRYQKDM 1125
Db 1141 ILWKLGFFKRYEKM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Umyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze

submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A:Cross-references: EMBL:Z29987; NID:G473098; PIDN:CAA82877.1; PID:G473099
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
P:169-344/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.0%; Score 1057; DB 2; Length 1178;

Best Local Similarity 27.8%; Pred. No. 3.2e-65;

Matches 341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;

Qy 1 FNLDTENAMTFQ--ENARGFGQSVVQL---QGSRVVVGAPQEIIVAAQNRGLSYOC--DYST 54
Db 27 YNVGLPKAKIFSPSSSEQFGYSVQQLTNPQGNWLLVGSWSPGPEPNRMDGVYKCPVDLPT 86
Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPOLLACQPTVHQTCSNTYVKGLC 107
Db 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLCTGCLPAHQCCNGOYATGIC 146
Qy 108 FLFGSNLRQPOQ---KPEALRCPCQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQL 164
Db 147 ----SDVSPDFQFLTFSFPAVQACPSL--VDVVVVCDESNSIYP--WEAVNKLVLKVFVTGL 199
Qy 165 K--KSKTLFSLMOYSEEFRIHFTFKBFQNNPNRSLIKPITQLLG--RTHTATGLRKVVRE 221
Db 200 DIGPKKTOVALIQYANEPRIIFNLNDFETKEDMVQATSETROHGGDLTNTFFRAIBFARDY 259
Qy 222 LFNITNGARKNAFKILFLTDGKFGDPLGYEDVIPLEDRGVIRY---VLGF--GDAFR 276
Db 260 AYQTSQGRPGATKVMVVVTDGESH--DGSKLKTVIOCCNDEILRFGIAVLGLNRLNALD 318
Qy 277 SEKSRQELNTVASKPRDRHVFQANNFEALTKVQNLREKIFAIEGTOTGSSSSFEHMSQ 336
Db 319 TKNLKIKETKAIAGTPTTERTYFFNVADBAALLEKAGTLEQIFSTEGTVQG--GDNFQWEMAQ 377
Qy 337 EGPSA--AITSNGPLSLTVSGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN--DAYLGYA 391
Db 378 VGSADYAPQNDILMLGAVGAPWMSGTLVQETSHKPVIFPKQAFDQVLDQRNHSFLGYS 437
Qy 392 AAIIILNRVOSLVGAPRYQHIGLVAMFRONTGWNESNANV-----KGTQIGAYFGASLCS 447
Db 438 VAAIISTEDGVHVFAGAPRANVTGQIVLYSVNK---QGNVTVIQSHRGDQIGSYFGSVLCS 494
Qy 448 VDVSNGSTDLVLIGAPHYVEQTR--GGQVSVCPPLPRGQARWQCDVAVLYGQGGQPWGRF 505
Db 495 VDVKOTITDVLVVGATPYMNDLKKEGKYLFTITKILNQHQ---FLEGPEGTGNARF 551
Qy 506 GAALTVLGDNVNGDKLTDVAIGAPGEDNKGAVYLFHGTSGSGISPSHSORIGASLSPR- 564
Db 552 GSAIAALSDINMDGDFNDVVGSPVENENSGAVIYNGHOGT--ITKYQSKILSNGAFRR 610
Qy 565 -LQYFGOSLGGQDLTMDGLVDLTGCAQGHVLLRQSVLRLKAIWEPNPREVARNVFE 623
Db 611 HLQFFGRSLDGYGDLNGDSITDVSIGALQGVQLWSQSIADVAIEALFTP----- 660
Qy 624 NDQVVKKEAGEVRVCLHVQKSTRDLREGQIQSVVYDIALD----SGRPHSRVAFNET 679
Db 661 -DKITLLNKDAKITLKLCFRAEFPAQONNOV---AILFNMTLDADGHSSRVSRTSGVFPREN 717
Qy 680 KNSTRROTQVGLGTQTCET--LKLQLPNCIEDPVSIVLRNLSFLVGTPLSAFGLNRLPVL 737
Db 718 SERFLQKNVNVQVKCSEHHISIQKPS---DVVNPLDLRVDISLENPQTS-----PAL 768
Qy 738 AEDAQRILTALFPPEKCNQNDNICQDDLSI-----TFSFMSLDCLVGGPREFNVTV 788
Db 769 EAYSETVKVFSIPFYKEGSDGICISDLILDVQQLPAIQTSF-----IVSNQNRLLTFS 823
Qy 789 VTVNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASST--EVSGALKST 847
Db 824 VILKRGESAYNTVLAEFSENLPF-----ASFSPMPVDGTEVTECEVSSQKSV 871
Qy 848 SCNHNHPIPEENSEVTNIFTVDVDSKASLGKLLKANVTSENMPRTNKTET--FQLELP 905
Db 872 TCDVGYPALKSEQQVTTTINFDNLQ--NLQNAAINFQAFSESO--ETNKADNSVSLTIP 928
Qy 906 VKYAVYVMTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQOR-----SLPISLV 955

```

Db 929 LLYDAEHLT-RSTNINFIYSSDENAPSVIK-----SVEDIGPKFIPFLSKVTAGSAPVMA 984
Qy 956 FLV-----PVRINQTVIMDRPQVTF-SENLS 980
Db 985 LVTHIQYTKENPLLYLTGIOTDQAGDISCTAEINPLKLPHTA-----PSVSFKNEFR 1040
Qy 981 STCHTKERLSHSDFLAELRKAPVNVGSIACVORIQCDIPFFGIQEBFNATLKNLSFDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMAEYFINVITRVNRT 1080
Qy 1041 YIKTSHNHLIVSTAEILFNDVSFTLPLPGOGAFVRSQTETKVEFPVNPPLPLIVGSSVG 1100
Db 1081 FASFTQVQLTAANEIDTHNPQFVIEENAVTIPLMIMKPTKAEVPT---GWIIGSIIA 1138
Qy 1101 GLLLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLLLAMTAGLWKLGFKQYKDM 1163

RESULT 10:
integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 chain
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 21-Jul-2000
A:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A:Reference number: A33998; MUID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <PAK>
A:Cross-references: GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A:Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R:Cattell, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIc, GPIIb and GPIIId)
A:Reference number: A56793; MUID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
A:Experimental source: platelet
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements
A:Reference number: A53117; MUID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, 'V', 18-21 <ZUT>
A:Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A:Note: authors translated the codon GTA for residue 17 as Leu
C:Genetics:
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:128031; OMIM:192974
A:Map position: 5q11.1-5q11.2
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 17.9%; Score 1054; DB 2; Length 1181;
Best Local Similarity 26.8%; Pred. No. 5.2e-65;
Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;

Qy 1 FNLDTENAMTFQ-ENARGCQSVVQL---OGRVVVCAPOEIVAAVNRGLYQC--DYST 54
Db 30 YNVLPEAKIFSGPSSEQFGYAVQQFINPKGNLLVGPSPWSPGPFENRMGDVYKCPVDLST 89

```

```

Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPPOLLACQPTVHQTCSENTYYKGLC 107
Db 90 ATCEKLNLTSTSIPIVNTBMTNMSLGLILTRNMGTFGLTCPLWAQCCNGQYTTTGV 149
Qy 108 FLFGSNLRQOPQKFPFALRGCEQSDSDIAFLIDGSGSIIIPHPRRMKELVSTIMEOLK-- 165
Db 150 SDISPDF-QLSASFSPATPCPSL-IDVVVWCDSENSIYPWD--AVKNFLEKPVQGLDIG 205
Qy 166 KSKTFLSLMQYSEEFRIHFTFKFQNNPNRSLIKPIQTQLLGRTHATATGLRKRKVVREL 224
Db 206 PKTQVGLIOYANNPRVFNLTNTYKTEEMIVATSTQYGGDLTNFCAIQVARYKVS 265
Qy 225 IYNGARKNAFKILFLTDCGKFGDPUGYEDVPELDEGVIRY---VLGF--GDAPRSEK 279
Db 266 AASGGRRSATKVMVVVTDGESH-DGSMLKAVIDQCNDHLIRFGIATVGLYLRNALDTKN 324
Qy 280 SRQELNTVASKPRDRHVFOANNFEALKTQVONLREKIFAIEGTQTQTCSSSEFHEMSQ 339
Db 325 LIKEIKAIASIPTEYFFNVSDAEALKEKAGTLGEQIFSEGTVOG-GDNFQEMSQVGF 383
Qy 340 SAAITSGNP--LLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGAAAI 394
Db 384 SADYSSQNDILMLGAVAGMGSTIVQKTSHGHLIFPKQAFDQILQDRNHSYVLGVSAA 443
Qy 395 ILNRVQSLVGLAPRYOHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLC 448
Db 444 ISTGESTHFVAGPRANYTGQIVLYSVN-----ENGNITVIAHRGDDGIGSYFGSVL 498
Qy 449 DVDSNGSTDLVLIGAPHYEOTR--GGQVSVCLPRGQARWQCDAVLGEQOQWGRFG 506
Db 499 DVDKOTITDVLVVGAPMYSDLKKEBGRVYLFTIKKGIHQH---FLEGPEGIENTRFG 555
Qy 507 AALTVLGVNVDKLTDAVIGAPCEEDNAGVYLFHGTSGSGISPSHSORIGS--KLSPR 564
Db 556 SAIAALSDINMGDFNDVIVGSPLENQNSGAVIYNHGQGT-IRTKYSKILGSDGAFRSH 614
Qy 565 LQVFGOSLGGQDLTMDGLVLTGVAQGHVLLRSQPLVLRVKAIMEFNPREVARNYFECN 624
Db 615 LQVFGRLDGYGDLNGDSITDSIGAFQVQVQLWSQSIADVAIEASTPEKI--TLVKN 672
Qy 625 DQVVVKEAGEVRVCLHVOKSTRDLREGIQSVVITYDLALD-----SGRPHSAVNETK 680
Db 673 AQII-----LKLCP---SAKFRPTKQNNQVAIVYNTILDADGFSRVTSRGLPKENN 721
Qy 681 NSRRTOVQLGLTQTC--ETLKLQPNCTEDPVSPIVLRNLNFSLVGTPLSAFGLNRPVLA 738
Db 722 ERCLQNMVNVQAQSCPEHIYIQEBS---DVNSLDLRLVDSLENPGTS-----PALE 772
Qy 739 EDAQLRFTALPFPKNCNDNICQDDLSITF-----SFMSLDCLVVGPPREFNVTVTRND 794
Db 773 AYSETAKVFSIPFKDCGEDGLCISDLVDVRQIPAAQEQPFIVSNQKRLTFSVTLKNK 832
Qy 795 GEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLACESAST-EVSGALKTSTCSINH 853
Db 833 RESAYNTGIVDFSENLF-----ASFSLPVDGTEVTCQVAASQKSVACDVGY 880
Qy 854 PIFPENSEVTENTFDVDSKASIGNKLLKANVTSENNMPRTNKTFFOLELPVKYAVYMW 913
Db 881 PALKREQQVTFITFNFDNLQ-NLQNASLUSFOALSQEBENKADNLVNLKIPILYDAEI- 938
Qy 914 VTSHGVTKYLVNFASENTSRVMQHOYVSNLQOR-----SLPISLVFLV----- 958
Db 939 ---HLTRSTININFEYISSDGNVPSIVHSFEDVGPKEIFSLKVTGTSVPVSMATVIHIP 995
Qy 959 -----PVRINQTVIMDRPQVTF--SENLSSTCHTKER 988
Db 996 YTKENKPLMYLTGVQTDKAGDISCNADINPLKIGQT---SSSVSPKSENF---HTKE- 1047
Qy 989 LSHSDFLAELRKAPVWVCSIAVCQRIQCDIPFFGIQEBFNATLKNLSFDWIKTSHNH 1048
Db 1048 -----LNCRTASCNSVTCLWKDVHMKGBYFVNVTTRINWNGTFASFTQT 1091

```

```

Db      618 QRIPSGGDKTKLFFGQSITHGMDLNGDGLTVITIGLGGAAALFWARDVAVVKVTWNFEP 677
Qy      614 REARVNVFECNDQWVGKEAG--EVRVCLHVQ-KSTRDLREQGIQSVWYDLDALDSGRP 670
Db      678 NKVNIQKNCR---VEGKETVCINATMCFHVKLKSKEDSIYEADLQ---YRVLDSLRLQ 730
Qy      671 HSRVAVNET-----KNSTRQTVGLGLTOTCETLKLQLPNCI-----BDPVSPIV 715
Db      731 ISRSFFSGTQERKIQRNITVRESE-----CIRHSFYMLDKDKHDFQDSVR 773
Qy      716 LRLNFSIVGTPLSAFGLNRLVLAEDAQRFTALFPEKQNGNDNICODDLSITFSFMSLD 775
Db      774 VTUDFNLT-DPENG-----PVLDDALPNSVHEHIFPAKOCGNKERCISDUTLNVSTTEKS 827
Qy      776 CLVVGPGPRE-FNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACES 834
Db      828 LLIVKSHQDKFNVLTVKNGDSAYNTRTVVQHSPLNIPSGIEIQKD-----SCES 879
Qy      835 ASSTEVSGALKSTCSINHIPIFENSEVTNITFDVDSKASLGNKLL-LKANTVSENMP 893
Db      880 N-----QNITCRVGPFLRAGETVTFKIIQFNTSHLSNAIHLSATSDSEEPLE 930
Qy      894 RTNKTFFQLELPVKYAV---YMWVTSHGYST-----KYLNTASENTSRVMQHOYQV 942
Db      931 SLNDNEVNISIPKYEVGLQFYSSAHEHISVAANETIPEFINST--EDIGNEINVEYTI 988
Qy      943 SNLGQSRSLP--ISLVF-----LVPVRLNQTWMD-----RP----- 971
Db      989 RKGGHPMPPELQISIFPNLTADGYPVLPIG-----WSSDNVNCRPSLEDPPGINS 1042
Qy      972 --QVTFs-----ENLSSTCHTKERLPSSHDFLAELRKAPVNVNCISIAVCORIQCDI 1019
Db      1043 GKXWITISKSEVLKRGTIQDCSSTC-----GVATITCSLLPSDLSQVNV 1085
Qy      1020 -----PFFGIQEEF---NATLKNLSFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQ 1070
Db      1086 SLLLWKPFP-IRAHFSLNLTLRGELK-----SENSLTLSSN----- 1123
Qy      1071 GAFVRSQTEKVPFEPVNPPLP--IVGSSVGGLLILALITAALYKLGPFKRYQKDWNSE 1128
Db      1124 ---RKRELAIQISKDGLPGRVPLWVLLSFAFAGLLMLLILALWKIGFPKRPPLKKQWEK 1180

RESULT 12
A41131
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N:Alternate names: Integrin alpha-4
C:Species: Mus musculus (house mouse)
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999
C:Accession: A41131; S16742
R:Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.
J. Cell Biol. 115, 1149-1158, 1991
A:Title: Cloning and expression of cDNAs for the alpha subunit of the murine ly
A:Reference number: A41131; MUID:92064645; PMID:1840602
A:Accession: A41131
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1039 <NU>
A:Cross-references: EMBL:X53176; NID:g51484; PIDN:CAA37316.1; PID:g51485
C:Superfamily: Integrin alpha-4 chain
C:Keywords: cytoskeleton; transmembrane protein

Query Match      11.3%; Score 666; DB 2; Length 1039;
Best Local Similarity 22.8%; Pred. No. 5.9e-38;
Matches 283; Conservative 183; Mismatches 405; Indels 368; Gaps 48;

Qy      1 FNLDTENAMTFQ-ENARGFQSVV-QLQGRS--VVVGAPOBIVAA--ORGSLYQCDY 52
Db      41 YNLDPENALLYQPSGTLFGYSVVLHSHGSKRWLLVGPATASWLSNVAVNPVGAIRCGI 100
Qy      53 STG---SCEPIRLQVP-----VEAVNMSLGLSLAATTSP-POLLACG----PTVHQ 95

```


Db 918 IV-----MLNT-EILKXSSVIO-----FMSRAKVVDPALRVYEAHGNPBEV 962
QY 1085 ---FEVFN---PLPLTVG---SSVGGLLLLALITAALYKLGFFKROYKDNM 1126
Db 963 TVVFEALNLEPRGYVGMIIAISLLVGLIFLLAVLLWKMGFFRRRYKEII 1015

RESULT 15
T31437
Integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.5% Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 2.4e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAAITSGNP--LLSTVGSVDWAGVFLYTSKEKSTFINNTRVDS----DMNDAYLG 389
Db 181 QAGFSGIIFDNSALVWAGPGSYLQGIYVQSLNRSV-VQATGESNTGTYSFDSNVRG 239
QY 390 YAAAI--ILNRVQSILVIGAPYQHI-GLVAMFRQNTGMWESNANVKGITQIGAYFGASLC 446
Db 240 YSLALGDFNGDVQDVVGTTPRAESLMGLVAIFDQNLQFN--QVMGTQIVAYFGYSVT 296
QY 447 SVDVDNSGTDLVLCAPHYEQTRGGQVSVCPFRQQRARWQCDVALYGEQ----- 498
Db 297 VVDI--NNDYDILLVGAFFMDDGPAIQ-----RWEAGAVYYLQNPDPVPGA 343
QY 499 -----GQPMGRFGAALTVLGVNGDKLTDVAIGAPCEDNRGAVYLFHTSGS 546
Db 344 SNRLSSTLIGGQTRSGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYHG-SAN 402
QY 547 GISPSHSQRIAGSKLS-PRLOYFGQSLSGQDLTMDGLVDLTVGAQ--GHVLLRSQPVL 603
Db 403 GLKSTPAQLTPSTLGHSGITTFGSLQGGQMDKNKYPDLLVGAESANTAVLIRTPV 462
QY 604 RUKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRLRGQIQSV----- 658
Db 463 SLDATLNTPE-----IGINLENTYE-LADGTWVTSFIAMT 497
QY 659 -----VTYDLALDSG-RPHSRVAFNETKSTRTQTVLGL-TQTCETLKQ 702
Db 498 CFTYTGNYLPDHDISYTVVDSGIIANRRAMPVNDMSEITKRRRLAVSTQFCPLRAY 557
QY 703 LPNCIEDPVSPVLRNLNPLSVGTPLSAFGN-----LRPVLAEADAQRLFTALFPPEK 753
Db 558 VGNISIEDKLTPIKVTLOYDL-----NNDESRLQPHILPIIDMATMTOTKQVSIQN 609
QY 754 NCGNDNICQDDLSITFSFMSLCLVVGGRPRENVTVVRNMGEDSYRTQVTFPPFLDLSY 813
Db 610 NCVN-NICIPDLQDVVT-PNLPNIVIGTQELTLDVSLNRRGEDAFQSSLSVYPLQF 667
QY 814 RKVSTLQNRQSQRWELACESASSTEVSGALKSTCSINHPIFPEN-----SEVTENIT 867
Db 668 VRL-----ERKANMDFSVTGSDD-----LRIITCDGNPMVGKNILEFGLTLSTFQVS 717
QY 868 FDVDS-----KASLGNKLLKANVTSENNMPTNKTEFQLELPVKYAVYVMVYVTSHGVTK 922

Search completed: November 25, 2003, 14:21:55
Job time : 19.9579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.17418 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNDLTENAMTFQENARGFGQ.....FKRQYKDMMSGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5823.5	99.1	1152	1	ITAM_HUMAN
2	4447	75.6	1153	1	ITAM_MOUSE
3	3446	58.6	1163	1	ITAX_MOUSE
4	3388	57.6	1162	1	ITAD_HUMAN
5	1536.5	26.1	1170	1	ITAL_HUMAN
6	1516.5	25.8	1163	1	ITAM_MOUSE
7	1136.5	19.3	1167	1	ITAE_MOUSE
8	1128	19.2	1179	1	ITAE_HUMAN
9	1085.5	18.5	1151	1	ITAI_HUMAN
10	1079.5	18.4	1189	1	ITAH_HUMAN
11	1071	18.2	1170	1	ITAB_BOVIN
12	1057	18.0	1178	1	ITAB_MOUSE
13	1054	17.9	1181	1	ITAB_HUMAN
14	1049	17.8	1180	1	ITAT_RAT
15	1042.5	17.7	1167	1	ITAG_HUMAN
16	666	11.3	1039	1	ITAA_MOUSE
17	642	10.9	1038	1	ITAA_HUMAN
18	630	10.7	1035	1	ITAA_XENLA
19	594.5	10.1	1032	1	ITAA_XENLA
20	571.5	9.7	1066	1	ITAB_CRISP
21	567.5	9.7	1053	1	ITAB_MOUSE
22	555.5	9.4	1053	1	ITAS_MOUSE
23	551.5	9.4	1034	1	ITAV_CHICK
24	542.5	9.2	1050	1	ITAA_XENLA
25	539	9.2	1044	1	ITAV_MOUSE
26	534.5	9.1	1130	1	ITAB_HUMAN
27	533.5	9.1	1072	1	ITAB_CHICK
28	532	9.0	1049	1	ITAS_HUMAN
29	531.5	9.0	1066	1	ITAB_HUMAN
30	530	9.0	1048	1	ITAV_HUMAN
31	514.5	8.8	1044	1	ITAB_CHICK
32	511.5	8.7	1091	1	ITAB_MOUSE
33	505	8.6	1179	1	ITAT_MOUSE

RESULT 1

ITAM_HUMAN	ID	ITAM_HUMAN	STANDARD	PRT	1152 AA
AC	P11215				
DT	01-JUL-1989	(Rel. 11, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)				
DE	(Neutrophil adherence receptor).				
GN	ITGAM OR CR3A OR CD11B.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88315033; PubMed=2457584;				
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;				
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";				
RL	J. Biol. Chem. 263:12403-12411(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88190151; PubMed=2833753;				
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;				
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88257215; PubMed=2454931;				
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;				
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";				
RL	J. Cell Biol. 106:2153-2158(1988).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=93123748; PubMed=8419480;				
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;				
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";				
RL	J. Immunol. 150:480-490(1993).				
RN	[5]				
RP	SEQUENCE OF 9-1153 FROM N.A.				
RX	MEDLINE=89098893; PubMed=2563162;				
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;				
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).				
RN	[6]				

ALIGNMENTS

RP SEQUENCE OF 1-9 FROM N.A.
 RX MEDLINE=92073318; PubMed=1583702;
 RA Shelley C.S., Arnaout M.A.;
 RT "The promoter of the CD11b gene directs myeloid-specific and
 developmentally regulated expression.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
 RN [7].

RP SEQUENCE OF 1-9 FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=92144986; PubMed=1346576;
 RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
 RT "Characterization of the myeloid-specific CD11b promoter.";
 RL Blood 79:865-870(1992).
 RN [8].

RP SEQUENCE OF 17-31.
 RX MEDLINE=87076671; PubMed=3539202;
 RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
 RT "N-terminal sequence of human leukocyte glycoprotein Mol;
 conservation across species and homology to platelet IIB/IIIA.";
 RL Biochim. Biophys. Acta 874:368-371(1986).
 RN [9].

RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
 RX MEDLINE=95171458; PubMed=7867070;
 RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
 RT "Crystal structure of the A domain from the alpha subunit of integrin
 CR3 (CD11b/CD18).";
 RL Cell 80:631-638(1995).
 RN [10].

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
 RX MEDLINE=96363671; PubMed=8747460;
 RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
 RT "Two conformations of the integrin A-domain (I-domain): a pathway for
 activation?";
 RL Structure 3:1333-1340(1995).
 RN [11].

RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
 RX MEDLINE=98362595; PubMed=9687375;
 RA Baldwin E.T., Saxer R.W., Bryant G.L. Jr., Curry K.A.,
 Fairbanks M.B., Finkel B.C., Garlick R.L., Heinrichson R.L.,
 Horton N.C., Kelley L.L., Wildner A.M., Moon J.B., Mott J.E.,
 Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
 RT "Cation binding to the integrin CD11b I domain and activation model
 assessment.";
 RL Structure 6:923-935(1998).
 RN [12].

RP 3D-STRUCTURE MODELING OF 17-616.
 RX MEDLINE=98226734; PubMed=9560195;
 RA Oxvig C., Springer T.A.;
 RT "Experimental support for a beta-propeller domain in integrin alpha-
 subunits and a calcium binding site on its lower surface.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
 CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
 ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
 AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
 IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
 THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
 PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
 FIBRINOGEN, FACTOR X AND ICAM-1. IT RECOGNIZES P1 AND P2 PEPTIDES
 OF FIBRINOGEN GAMMA CHAIN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
 ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
 GRANULOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide Cdb1b entry;
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdb1b.htm".
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; J03925; AAA59544.1; -;
 DR EMBL; M18044; AAA59491.1; -;
 DR EMBL; J04145; AAA59903.1; -;
 DR EMBL; S52227; AAB24821.1; -;
 DR EMBL; S52152; AAB24821.1; JOINED.
 DR EMBL; S52153; AAB24821.1; JOINED.
 DR EMBL; S52154; AAB24821.1; JOINED.
 DR EMBL; S52155; AAB24821.1; JOINED.
 DR EMBL; S52157; AAB24821.1; JOINED.
 DR EMBL; S52159; AAB24821.1; JOINED.
 DR EMBL; S52161; AAB24821.1; JOINED.
 DR EMBL; S52164; AAB24821.1; JOINED.
 DR EMBL; S52165; AAB24821.1; JOINED.
 DR EMBL; S52167; AAB24821.1; JOINED.
 DR EMBL; S52169; AAB24821.1; JOINED.
 DR EMBL; S52170; AAB24821.1; JOINED.
 DR EMBL; S52173; AAB24821.1; JOINED.
 DR EMBL; S52174; AAB24821.1; JOINED.
 DR EMBL; S52180; AAB24821.1; JOINED.
 DR EMBL; S52181; AAB24821.1; JOINED.
 DR EMBL; S52184; AAB24821.1; JOINED.
 DR EMBL; S52189; AAB24821.1; JOINED.
 DR EMBL; S52191; AAB24821.1; JOINED.
 DR EMBL; S52192; AAB24821.1; JOINED.
 DR EMBL; S52202; AAB24821.1; JOINED.
 DR EMBL; S52212; AAB24821.1; JOINED.
 DR EMBL; S52213; AAB24821.1; JOINED.
 DR EMBL; S52216; AAB24821.1; JOINED.
 DR EMBL; S52219; AAB24821.1; JOINED.
 DR EMBL; S52220; AAB24821.1; JOINED.
 DR EMBL; S52221; AAB24821.1; JOINED.
 DR EMBL; S52222; AAB24821.1; JOINED.
 DR EMBL; S52226; AAB24821.1; JOINED.
 DR EMBL; M76724; AAA58410.1; -;
 DR EMBL; M84477; AAA51960.1; -;
 DR PIR; A31108; RWHU1B.
 DR PDB; 1A8X; 17-JUN-98.
 DR PDB; 1BHO; 18-NOV-98.
 DR PDB; 1BHQ; 18-NOV-98.
 DR PDB; 1IDN; 25-NOV-98.
 DR PDB; 1IDO; 01-AUG-96.
 DR PDB; 1JLM; 11-JAN-97.
 DR PDB; 1M1U; 07-AUG-02.
 DR Genew; HGNC:6149; ITGAM.
 DR MIM; 120980; -;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWF_1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VWFA; 1.
 DR Signal; 3D-structure; Repeat; Magnesium; Calcium.
 KW Signal; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 FT SIGNAL 1 16
 FT CHAIN 17 1152 INTEGRIN ALPHA-M.
 FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1105 1128 POTENTIAL.
 FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 31 84 FG-GAP 1.

Query Match		99.1%;	Score 5823.5;	DB 1;	Length 1152;		
Best Local Similarity		98.8%;	Pred. No. 0;				
Matches 1123;		Conservative	8;	Mismatches	5;	Indels	1; Gaps 1;
QY	1	FNLDTENAMTFOENARFGQSVVQLOGSRVVVGAPOEIVAAANORGSLYOCYSTGSCPEI	60				
DB	17	FNLDTENAMTFOENARFGQSVVQLOGSRVVVGAPOEIVAAANORGSLYOCYSTGSCPEI	76				
QY	61	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK	120				
DB	77	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK	136				
QY	121	FPALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMQVSEBF	180				
DB	137	FPALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMQVSEBF	196				
QY	181	RHFTPEKQNNPNSLKPITOLGRTHATGLRVKVVRELFINITNGARKNAFKILFL	240				
DB	197	RHFTPEKQNNPNSLKPITOLGRTHATGLRVKVVRELFINITNGARKNAFKILVVI	256				
QY	241	TGGEKFGDPLGYEDVPEADREGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHVFOAN	300				
DB	257	TGGEKFGDPLGYEDVPEADREGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHVFOAN	316				
QY	301	NPEALKTVONQLEKIFATGEGTQSGSSPFHEMSQEGFSAATISNGPLSTVGSYDMAG	360				
DB	317	NPEALKTVONQLEKIFATGEGTQSGSSPFHEMSQEGFSAATISNGPLSTVGSYDMAG	376				
QY	361	GVFLYTSKSKSTFINNTRVDSNDNDAYLGAAAAIILNRVQSLVLCAPRYOHTGLVAMFR	420				
DB	377	GVFLYTSKSKSTFINNTRVDSNDNDAYLGAAAAIILNRVQSLVLCAPRYOHTGLVAMFR	436				
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCPL	480				
DB	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCPL	496				
QY	481	PRQORARWOCDAVLYGEQOPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAYVLF	540				
DB	497	PRG-RARWOCDAVLYGEQOPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAYVLF	555				
QY	541	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ	600				
DB	556	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ	615				
QY	601	PVLVRKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHWQKSTRDLREGQIOSVVT	660				
DB	616	PVLVRKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHWQKSTRDLREGQIOSVVT	675				
QY	661	YDLALDSGRPHSRAVFNENKSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF	720				
DB	676	YDLALDSGRPHSRAVFNENKSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF	735				
QY	721	SLVGTPLSAFNGLRPVLAEDAQRLFTALPPFKKNCNDNICQDDLITFSFMSLDCLVVG	780				
DB	736	SLVGTPLSAFNGLRPVLAEDAQRLFTALPPFKKNCNDNICQDDLITFSFMSLDCLVVG	795				
QY	781	GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLONORSQSRMLACESASSTEV	840				
DB	796	GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLONORSQSRMLACESASSTEV	855				
QY	841	SGALKSTCSINHPIEPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF	900				
DB	856	SGALKSTCSINHPIEPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF	915				
QY	901	QLELPVKYAVYMWVTSKGYSTKYNFTASNTSRVMOHQYQVSNLQORSPLISLVLFPV	960				
DB	916	QLELPVKYAVYMWVTSKGYSTKYNFTASNTSRVMOHQYQVSNLQORSPLISLVLFPV	975				
QY	961	RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCIP	1020				
DB	976	RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCIP	1035				

1021 PFGIOBERNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFVTLPLPGQCAFVRSOTET 1080
1036 PFGIOBERNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFVTLPLPGQCAFVRSOTET 1095
1081 KVEPEVENPLPLIVGSSVGLLLALITAAALYKLGFFKROVKDMMSGGPPGAEPQ 1137
1096 KVEPEVENPLPLIVGSSVGLLLALITAAALYKLGFFKROVKDMMSGGPPGAEPQ 1152

RESULT 2
ITAM MOUSE
ID ITAM MOUSE STANDARD; PRT; 1153 AA.
AC P05555;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1). ITGAM.
GN ITGAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88312584; PubMed=3044779;
RA Pytela R.;
RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";
RL EMBO J. 7:1371-1378(1988).
RN [2]
RP SEQUENCE OF 11-45 FROM N.A.
RX MEDLINE=86287312; PubMed=2942940;
RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;
RT "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
RN [3]
RP SEQUENCE OF 17-28.
RX MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3b. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION. IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the ENBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL; X07640; CAA30479.1; .
EMBL; M14293; AAB39484.1; .
PIR; S00551; S00551.
HSP; P11215; IABX.
MG; MG1:96607; Itgam.
InterPro: IPR000413; Integrin alpha.

Qy	61	RLQVPVEAVNMSLGLSLAATTSPQILLACQPTVHQTCSENTYYKGLCFLPGSNLRQOPQK	120
Db	77	PLQVPPPEAVNMSLGLSLAVSTVPQILLACQPTVHQCKENTYYNGLCYLFPGSNLLRPPQQ	136
Qy	121	FPFALRCPOEDSDIAFLIDGSGIIPHDFRRMKEIIVSTIMEOLKKSKTLLFSLMQYSEEF	180
Db	137	FPFALRECPQESDIIVFLIDGSGIINIDQKWKFPVSTWBOFKKSKTLLFSLMQYSDFF	196
Qy	181	RIHFTFKFQNNPNRSLIKPIITOLLGRTHATGLRKRVKRELFINIINGAKNAFKILL	240
Db	197	RIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTASGIRKRVKRELPHKTINGARENAKILL	256
Qy	241	TDEKFGDPLGYEDVITPELDRBGVIRVVLGFGDAPRSEKSRQELNNTVASKPPRDHVP	300
Db	257	TDEKFGDPLGYKDVITPEADRAGVIRVIVGVNAFNKPPSRRELDITIASPAGEHVP	316
Qy	301	NFEALKTQVQLREKIFAIBEGTQTGSSSSFEHEMSQEGFSAATISNGPLLISTVGSVD	360
Db	317	NFEALNTIQLOKEIFAIBEGTQTGSSSSFEHEMSQEGFSAITISNGPLLGSGVSPD	376
Qy	361	GVFLYTSKEKSTFINMTRVDSDMNDAYLGYAAAILLRNVQSLVLGAPRVOHIGLVAM	420
Db	377	GAFLYTSKDKVTFINTTRVDSDMNDAYLGYASAVILLRNVQSLVLGAPRYQHIGLV	436
Qy	421	QNTGMWESNANVXGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGQVSV	480
Db	437	ENFGTWEPHTSIYKGSQIGSYFGASLCSVDMDADGNTNLIIGAPHYYEYTRGGQV	496
Qy	481	PRGORARWQCDVLYIGSQQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAY	540
Db	497	PRG-RARWQCEALLHGQGHQPMWRFGAALTVLGDVNGDKLTDVAIGAPGQENQGA	555
Qy	541	HGTSGSGISPSHSORISAGSKLSRLOYFGOSLGGQDLTMDGLVDLTVGAQGHVLL	600
Db	556	YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGKDLTMDGLMDLAVGAQGHLL	615
Qy	601	PVLRVKAIMBFNPREVARNYFECNQDVYVKEAGEVRVCLHVOKSTRDRLEGOIQSV	660
Db	616	PVLRLAETMEFSPKKVARSVFACQEVQLKNKDAGEVRVCLVRKNTKDRLEGGDIQ	675
Qy	661	YDLALDSGRHSRAVFNETHKNSTRQTOVLGLTQCTETIKLOLPNCIEDVSPVLPL	720
Db	676	YDLALDVRGRIRAFFDETKNNTRRRTQVGLMQKCTKLUIIPDCVDDSVSPILRL	735
Qy	721	SLVCTPLSAFONLRPVLAEDAQRULFTALFPFKNCGNDNICQDDLSITFSFMSLD	780
Db	736	TLVGEPLRSFGNLRPVLAMDARFTTAMFPFKNCGNDNICQDDLSITMSAMGLD	795
Qy	781	GPREFNVTVVRNDEGSYRQTVTFPPLDLVSRKVSTLQNRQSQRWRL-ACESAS	839
Db	796	GPQDFNMSVTLNRNDEGSYGTQTVVYYPGSLVSRKDSQSONPLTKPWFVKPAES	855
Qy	840	VSGALKSTCSINHPILPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTN	899
Db	856	GHGALKSTTWINHPILPANSEVTFNITFDVDSHASFONKLLKAIIVASENNMGR	915
Qy	900	FQLELPVKIAYMYVVTSHGVSTKYLNFTASENTSRRVMOHQYQVSNLQORSUP	959
Db	916	FQLELPVKIAYMYVTSDESSIRYLNFTASEMTSKVIHQYQFNNLQORSUPSV	975
Qy	960	VRLNQTVIWDPRQVTFSENISLSTCHTERLUPSHSDFLAELRKAPVNVCSIA	1019
Db	976	VQINNVTVWDHPQVIFSONUSSACHTEKSPPHNFRDQERTPVLNCSVAVCKRI	1035
Qy	1020	PFFGIQBEFNATLKNLSFDWYIKTSHNNHLLIVSTAILFNDSVFTLLPGQAF	1079
Db	1036	PSFNTQELFNVTLKNLSFDWYIKTSHGHLLVSTTEILFNDSAFALLPGQES	1095
Qy	1080	TKVBPFVEVNPDLPLIVSSVCGLLLLLALITAALYKGLFFKQYKOMMSEGGP	1137
Db	1096	TKVBPEYVHNPPVLIIVSSIGLVLLALITAGLYKGLFFKQYKOMMNEAAPQ	1153

RESULT 3

ITAX_HUMAN STANDARD; PRT; 1163 AA.

AC P20702;

DT 01-FEB-1991 (Rel. 17, Last Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95

DE alpha chain) (leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).

GN ITGAX OR CD11C.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88166645; PubMed=3327687;

RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;

RT "cDNA cloning and complete primary structure of the alpha subunit of

RT a leukocyte adhesion glycoprotein, p150,95.";

RL EMBO J. 6:4023-4028(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=90153906; PubMed=2303426;

RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;

RT "Genomic structure of an integrin alpha subunit, the leukocyte

RT p150,95 molecule.";

RL J. Biol. Chem. 265:2782-2788(1990).

RN [3]

RP ERRATUM.

RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;

RL J. Biol. Chem. 265:12750-12751(1990).

RN [4]

RP SEQUENCE OF 20-43.

RX MEDLINE=87167596; PubMed=3549901;

RA Miller L.J., Wiebe M., Springer T.A.;

RT "Purification and alpha subunit N-terminal sequences of human Mac-1

RT and p150,95 leukocyte adhesion proteins.";

RL J. Immunol. 138:2381-2383(1987).

CC -1- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT

CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL

CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY

CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X

CC ASSOCIATES WITH BETA-2.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND

CC GRANULOCYTES.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: Contains 1 VWFA domain.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC -1- DATABASE: NAMES=PROW; NOTE=CD guide CD11c entry;

CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC or send an email to license@isb-sib.ch.

CC -----

DR EMBL; M81695; AAA5180.1; -

DR EMBL; Y00093; CAA68283.1; -

DR EMBL; M29165; -: NOT ANNOTATED CDS.

DR EMBL; M29487; AAA51620.1; ALT SEQ.

DR EMBL; M29482; AAA51620.1; JOINED.

DR EMBL; M29483; AAA51620.1; JOINED.

DR EMBL; M29484; AAA51620.1; JOINED.

DR EMBL; M29485; AAA51620.1; JOINED.

DR EMBL; M29486; AAA51620.1; JOINED.

DR PIR; A36584; RWHUIC.

DR PDB; 1N3Y; 18-FEB-03.

DR Genew; HGNC:6152; ITGAX.

DR MIM; 151510; -

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.

DR GO; GO:0004872; F:receptor activity; TAS.

DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PRO1185; INTEGRINA.

DR PRINTS; PRO0453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 5.

DR SMART; SM00327; VWA_1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS50234; VWFA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

KW Signal; Magnesium; Calcium; Repeat; 3D-structure.

FT SIGNAL 1 19

FT CHAIN 20 1163 INTEGRIN ALPHA-X.

FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1108 1128 POTENTIAL.

FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).

FT REPEAT 34 87 FG-GAP 1.

FT REPEAT ? ? FG-GAP 2.

FT DOMAIN 165 351 VWFA.

FT REPEAT ? ? FG-GAP 3.

FT REPEAT 402 453 FG-GAP 4.

FT REPEAT 455 517 FG-GAP 5.

FT REPEAT 518 576 FG-GAP 6.

FT REPEAT 581 633 FG-GAP 7.

FT CA_BIND 466 474 POTENTIAL.

FT CA_BIND 530 538 POTENTIAL.

FT CA_BIND 593 601 POTENTIAL.

FT SITE 1131 1135 GPFKR MOTIF.

FT DISULFID 69 76 BY SIMILARITY.

FT DISULFID 108 126 BY SIMILARITY.

FT DISULFID 655 712 BY SIMILARITY.

FT DISULFID 771 777 BY SIMILARITY.

FT DISULFID 848 863 BY SIMILARITY.

FT DISULFID 998 1022 BY SIMILARITY.

FT DISULFID 1027 1032 BY SIMILARITY.

FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 490 490 G -> A (IN REF. 2).

FT CONFLICT 756 756 L -> D (IN REF. 2).

SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;

Query Match 58.6%; Score 3446; DB 1; Length 1163;

Best Local Similarity 60.5%; Pred. No. 1.5e-227;

Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGVAQEI VAAHQGSLYOCDYSTGSCPEI 60

DB 20 FNLDTBELTAFRVDSAGFGSDSVQVANSVWVGAPQKITAANQTGLGYCGYSTGACEPI 79

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHTCSEYTYVKGICFLFGSLNLRQOPK 120

DB 80 GLQVPEAVNMSLGLSLASTTSPLLACGPTVHHECGRMVLTGLCFLLGPT--QLTOR 137

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKKSKTFLFSLMQVSEF 180

DB 138 LPVSRQECPCRBQDQIVFLIDGSGSISSRNFMFVRAVISQFORPSTQFSLMQFSNKF 197

DR EMBL; U40274; AAB60634.1; .
DR EMBL; U40275; AAB60635.1; .
DR EMBL; U40276; AAB60636.1; .
DR EMBL; U40277; AAB60637.1; .
DR EMBL; U40278; AAB60638.1; .
DR EMBL; U40279; AAB60639.1; .
DR EMBL; U40278; AAB60638.1; JOINED.
DR EMBL; AF187881; AAF62875.1; .
DR HSSP; P11215; IABX.
DR Genew; HGNC:6146; ITGAD.
DR MIM; 602453; .
DR GO; GO:0008305; C-integrin complex; TAS.
DR GO; GO:0004895; P:cell adhesion receptor activity; TAS.
DR GO; GO:0016337; P:cell-cell adhesion; NAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium;
KW Magnesium.
FT SIGNAL 1 17
FT CHAIN 18 1162
FT DOMAIN 18 1100
FT DOMAIN 1101 1124
FT DOMAIN 1125 1162
FT REPEAT 32 85
FT REPEAT ? ?
FT DOMAIN 150 332
FT REPEAT 350 400
FT REPEAT 401 452
FT REPEAT 454 516
FT REPEAT 518 576
FT REPEAT 581 633
FT CA_BIND 465 473
FT CA_BIND 530 538
FT CA_BIND 593 601
FT SITE 1127 1131
FT DISULFID 167 74
FT DISULFID 106 124
FT DISULFID 655 710
FT DISULFID 769 775
FT DISULFID 846 861
FT DISULFID 994 1018
FT DISULFID 1023 1028
FT CARBOHYD 59 59
FT CARBOHYD 87 87
FT CARBOHYD 99 99
FT CARBOHYD 391 391
FT CARBOHYD 691 691
FT CARBOHYD 733 733
FT CARBOHYD 873 873
FT CARBOHYD 957 957
FT CARBOHYD 1046 1046
FT CONFLICT 500 500
FT CONFLICT 515 518
FT CONFLICT 825 825
FT CONFLICT 984 984
SQ SEQUENCE 1162 AA; 126885 MW; F296A1A35455D77D CRC64;
Query Match 57.6%; Score 3388; DB 1; Length 1162;
Best Local Similarity 58.9%; Pred. No. 1.4e-223;
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;
1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAORGSLYQCDYSTGSCPEI 60

QY


```

FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 730 730 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 862 862 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 885 885 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 897 897 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1060 1060 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC 954 954 Q -> QGVHGLVEMOTSKOILCRPAGDAEHTVGAQGEPLCP
/FTIGVSP 002738.
R -> W (IN REF. 1 AND 2) .
Y -> I (IN REF. 2) .

214 214
660 660
155 155
164 164
165 166
169 185
186 188
191 198
202 206
208 213
217 221
222 223
229 229
233 243
244 244
247 249
250 250
253 254
256 263
274 276
277 278
280 286
288 290
293 297
298 299
300 302
307 310
311 314
318 318
319 328
329 330
SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match 26.1%; Score 1536.5; DB 1; Length 1170;
Best Local Similarity 34.2%; Pred. No. 6.9e-97;
Matches 401; Conservative 209; Mismatches 463; Indels 101; Gaps 37;

QY 1 ENLDTENAMTFQ--ENARFGOSVVOLOQSRVVVGAPQEIIVANORGSLYQCDYSTGSC 58
Db 26 YNLVDVGARSFSPRAGRHFVGYKVLQV-GNGVTVGAPGE---GNSGSLYQCOSGTGHL 81
QY 59 PIRLOVPEAVNMSLGLSLAATTPPQLLACGPTVHTCSENTYVKGCLFLFGSNLR--- 115
Db 82 PVTLR-GSNYTSKYLGTATDPTDGSILACDPLGSLRTCDQNTYLSGLCYLFRQNLQGP 140
QY 116 -QOPQKFPALRCPOEDSDIAPLDGSGSIIPHDPRMKELVSTIMEOLKXSKTLFSLM 174
Db 141 LQGRPGFOECIKG---NVDLVFLFDGMSLQDPEFKILDFMKDVKMKLSNTSYQFAAV 196
QY 175 OYSEEFRIHTFKEFQNNPRSLIKPIITQLLGRTHATCLRVVPELFININGARKNAF 234
Db 197 QESTSYKTFEFDSDYKVKRQDPDALLKHVKHMLLTNTFGAINVATEVFREELGARPDA 256
QY 235 KILFLLTDEGKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSRQELNTVASKPPRD 294
Db 257 KVLIIITDGE--ATDSGNDAKD-----IIRVIIGKHFKQKESQETLHKFPASKPASE 309
QY 295 HVFQANNFALKTVQNLREKIFAIEGTQTGSSSFEHEMSQBGSAATISNGPLISTVG 354

```

```

310 FVKILDTFEKLDLFTLELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRSHAVVAVG 369
355 SYDMAGGVF-LYTSKEKSTFFINMTVDSMDNDAYLGYAAA-IILRRVQSLVGLGAPVQH 412
370 AKDMAGGVLKADLQDDTFIGNEPLTPEVRAGYLGTVTWLPSRKQTSLLASGAPRYQH 429
413 IGLVAMFR--QNTGMWESNANVKTQICAGYFASLCSVDVDSNGSTDLVLGAPHYEQT 470
430 MGRVLLFQSPQGGHMSQVQTIHGTFQISYFGEELCGVDVQDGETELLIGAPLYFEQ 489
471 RGGQVSVCPPLPRGQARWQCDAV--LYGEOGPMGRFGAALTVLGDVNGDKLTDVAIGAP 528
490 RGRGVFIY-----QRRQLGFVEVSLQDGPGLGRFGEAITALTDINGDGLVDVAVGAP 544
529 GEEDNRGAVYLPHTGSGISPSHQSRIAGSKSLPRLQYFGQSLSGGQDLTMDGLVDLT 588
545 LEE--QGAVYIFNGRHG-GLSPQPSQRIEGTQVLSGIQWFGKRSIHGVNDLEGDLAVAV 601
589 GAQGHVLLRSQPLRVKAIMFNPREVARNVFENDQVV-KGKAGEVRVCLHVQKSTR 647
602 GAESQMIVLSSRPVDMVTLSFSPAETPVHEVECSYSTSNMKKEGVNITICFQI-KSLY 660
648 DRLREGIOQSVVTVYDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQPCNI 707
661 PQF-QGRIVANLTYTLQDGHRTRRGLFPGGRHRLRNIAVT-TSMCTDIFSFPVPCV 718
708 EDPVSPVILRLNFSL---VGTPLS--AFGN-----LRPVLABDAORLFTALPFFKNCN 757
719 QDLISPINVSLNFWLWEEBEGTPDQRAQKQDIPILRPSLHSETWEI-----PFKNCE 773
758 DNICODLSITPSFMSLCLVVGPRFNVTVTVRNDGEDSYRTQVTFPPFLDLVSRKVS 817
774 DKKCEANLRVSPSPARSRALRTAPASLSVELSLNLEEDAYWVQLDLHFPPLGSPRKVE 833
818 TLQORSQSRWPLACES--ASTEVSGLKSTSCSINHPIPFENSEVTNITFDVDSKAS 875
834 ML---KPHSQIPVSCSELPEESRLLSRAL---SCNVSSPIFKAGHSVALOMMENTLVNS 887
876 LGNKLLKANVTSENN---MPRTNKTBFQLELPVKYAVYVVTSHGVSSTKYNFTASEN 931
888 WGSVELHANVTNNEDSDLEDNSATTI---IPILYPINILIQOEDSTLVSVFPPKGP 944
932 TSVVMQHQYQV---SNLQORSIP-LSLVFLVPLVRLNQTVWDRPQVTFSENLSTCHTK- 986
945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVGVGPQPPSEGPITHQWSVQMEPPV--PCHYED 1002
987 -RLPSSHSD--FLAELRKAPVVCISIAVCQRQCQDIPFGIGQEEFNATLKGNSLSPDWYIK 1043
1003 LERLPDAAEPCLPGALFRCPVV-----FROEILVQVIGTLELVGEIE 1044
1044 TSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTETKVEPPEVFNPLPLIVGSSVGGLL 1103
1045 AS-SMPSLCSLSISFNSKPHLYGSNASL-AQVVMKVVDVYVEKQMLYLVLSGIGLL 1102
1104 LLALITAAALYKLGFPKQYKQKMMSEG-GPPGAP 1136
1103 LLLLIFVLYKGVFKRNLKEKMEAGRVGPNIP 1136

RESULT 6
ITAL MOUSE
ID ITAL MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein IFA-1
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)
DE (CD11a).
GN ITGAL OR IFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91268576; PubMed=2051027;
RA Kaufmann Y., Tseng E., Springer T.A.;
RT "Cloning of the murine lymphocyte function-associated molecule-1
RL alpha-subunit and its expression in COS cells.";
RJ J. Immunol. 147:369-374(1991).
RN [2]
RP SEQUENCE OF 24-42.
RX MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RL glycoproteins and unexpected relation to leukocyte interferon.";
RN Nature 314:540-542(1985).
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
CC ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
CC AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
CC SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
CC LEUKOCYTES RECRUITMENT.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWF DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWF domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60778; AAA39426.1;
DR PIR; I56126; I56126.
DR HSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VFWA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium;
KW Repeat.
FT SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT ? ?
FT DOMAIN 148 334
FT REPEAT ? ?
FT REPEAT 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT CA_BIND 466 474
FT CA_BIND 528 536
```

```
FT CA_BIND 588 596
FT SITE 1111 1115
FT DISULFID 70 77
FT DISULFID 108 126
FT DISULFID 147 199
FT DISULFID 651 705
FT DISULFID 767 773
FT DISULFID 840 856
FT DISULFID 993 1009
FT DISULFID 1017 1048
FT CARBOHYD 86 86
FT CARBOHYD 185 185
FT CARBOHYD 270 270
FT CARBOHYD 444 444
FT CARBOHYD 668 668
FT CARBOHYD 696 696
FT CARBOHYD 724 724
FT CARBOHYD 728 728
FT CARBOHYD 776 776
FT CARBOHYD 857 857
FT CARBOHYD 880 880
FT CARBOHYD 890 890
FT CARBOHYD 899 899
FT CARBOHYD 927 927
FT CARBOHYD 1056 1056
SQ SEQUENCE 1163 AA; 128343 MW; A7A3078489E8232F CRC64;

Query Match 25.8%; Score 1516.5; DB 1; Length 1163;
Best Local Similarity 33.9%; Pred. No. 1.6e-95;
Matches 396; Conservative 216; Mismatches 466; Indels 89; Gaps 36;

Qy 1 FNLDTENAMTFQENA-RGFGQSVVQLGSRVAVVGAPOEIVAAQNRGSLYQDSTGSCPEP 59
Db 24 YNLDTREPTQSLAQAGRHFGVQLQIEDG-VVVGAPCE---GDNTGGLYHCRSTSEFCQP 79
Qy 60 IRLOVPVEAVNMSLGLSLAATSPQQLACGPTVHTQCSENTYVYKGLCFPLGSLNRQPO 119
Db 80 VSLH-GSNHTSKYLGMLTADAAKGLSLACDPLGSLRTCDQNTYLSGLCYLFPPOSLEPML 138
Qy 120 KPPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKSLFSLMQYSEE 179
Db 139 QNRPAYOECKMGKVDLVFLFDGSGSLDRKDFEKLFEKMDVMKLSNTSVQFAAVQSTD 198
Qy 180 FRIHFTPEKF-QNNPNPRSLIKPTQLLGRTHRTATGLRKVYVRELNFNTNGAKNAFKILF 238
Db 199 CRTEFTFLDYVKQNKNPVLLGSVQPMFLTNTFRAINVYVAHVKEESGARPDATKVLV 258
Qy 239 LLTDGEKF--GDPLGYEDVIPELDREGVIRVYVLFGFGDAFSEKSRQELNVTASKPPRDHV 296
Db 259 IITDGEASDKGNISAAHD-----ITRYIIGIKHFVSQKQKTLHFASEPVEEFV 309
Qy 297 FOANNFEALKTQVQNLREKIFATEGTQTGSSSSSEHEMSOEGPSAAITNSGPLLSTVGSY 356
Db 310 KILDTPEKLKDLFTDLORRIYAIETNRQDLTSFNWELSSGSIADLSKGHAVVGAVGAK 369
Qy 357 DWAGGVF-LYTSKEKSTFFINMTRVDSMDNDAYLGAAAA-IILNRVQSLVGLGAPRYQHIG 414
Db 370 DWAGGFLDLREDLQATFVGQEPVLTSDVRGCGYLGTVVAMWTSRSSRLAAGAPRYQHV 429
Qy 415 LVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRG 472
Db 430 QVLLFQAPAEAGGRWNQTKIEGTQIGSYFGELCSVDLDQDGEAEELIIGAPLFFGQRG 489
Qy 473 GOVSVCPPLPRGORARWOCDAVLVGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPCEED 532
Db 490 GRVFTY---QRRQLFEWVSELOQDPGYPLGRFGAALTALTDINGDRLTDVAVGAPLEE- 545
Qy 533 NRGAVYLFHGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGODLTMGDLVLTVGAQG 592
Db 546 -QGAVYIFNGKPG-GLSPQPSQRIQGAQVFPFGIRWFGRSIHGVKDLGDLADVVVGAEG 603
Qy 593 HVLLLRQPVLRVKVAMEFNPREVARNVFCNDQVVGKQKAG-EVRVCLHVQKSTRDLRLR 651
```


FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128993 MW; B8331C115DCCCFD CRC64;
Query Match 19.3%; Score 1136.5; DB 1; Length 1167;
Best Local Similarity 28.6%; Pred. No. 1.6e-69;
Matches 353; Conservative 213; Mismatches 462; Indels 205; Gaps 43;
QY 1 FNLOTENA--WTFQENARGFGQSVVLOGSRVVVGAQPEIVAAHQGS-----LYOCDY 52
DB 20 FNMVDWAWTALQFGAPAVLSSLLHLDPSN-----NOTCLLVARRSSNRNTAALYRCAL 74
QY 53 STGCEPIRLQVPEAVNMSLGLSLAATT--SPQLLAC-GPTVHQTCSNTYVYKGLCFL 109
DB 75 SI-SPDEIACQ-PVEHICMPKRGVGVTLVGNHNGVLVCIQVQARKFRSLNSELTGACSL 132
QY 110 FGSNLROQPKFPKALRG-----C-----PQE 131
DB 133 LTPNLDLQAQAYFSDLEGLDPGAHVDSGDYCRSGKSGTGBEEKSARRRRRTVEDEED 192
QY 132 DSDIAFLDGSIIIPHDPRMKELVSTIMEQL--KSKTFLSLMOYSEEFRIHFTPEF 189
DB 193 GTEIAVLDSGSGSPDFQAKNFISTMMNFVEKFCFECFALVQVGAIVTQTFDLDQES 252
QY 190 QNNPNRSLIKPITQLLCRTHRTATGLRKVRELNITNGARKNAFKILFLITDGEKFGDP 249
DB 253 RDINASLAKVSIQVQKEVTTASMQHVLNDIIFPSRGSRKALKVMVLTDCGIFGDP 312
QY 250 LGYEDVPELDREGVIRYVLGFGDAFRSEKSRBELNTVASKPRDHVFPQANFALKTVQ 309
DB 313 LNLTVINSKMGVVRFAIGVGRFKNNTYRELKLIASDPKEAHTPKVTNYSGALDGL 372
QY 310 NQLEKPIAIGTQTGSSSEFEHMSQEGFAAITSNCP-LLSTVGSVDWAGGVFLY-TS 367
DB 373 SKLQORIYHMEGT---VDALQYQLAQTGFSQAQILDKGQVLLGTGAPFNWGGALLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DNMDAYGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFRONTGM 425
DB 430 NGRGCFNLQTAKEUSRTQYSLGSLAVLHKANGISVAGAPRHLKAGAVFELKEDR- 488
QY 426 WESNA---NVKGTQIGAYFGASLCSDVDNSGSTDVLVIGAPHYYEOTRGQGVVCPILR 482
DB 489 -EEDAFVRIEAGEQMGVFGVLCVPVDIDMDGTTDFLLVAAPFFVHIRGEEGRVYVQPE 547
QY 483 QARARWQCDVLYGEOGPWGRFGAALTVLGVNQCGLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSGHPGLTNSRFGPAMAAGVDINDQKFTDVAIGAPLEFGAGGASVGS 606
QY 537 VYLFHGTSGSISPSHSQRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGAGQHVL 596
DB 607 VYINGHSG-GLYDPSQIRASSVASGLHYFGMSVSGLDENGDLADIIVGSRDSAVV 665
QY 537 LRSQVLVRKAIEMFNPEARVARNFECNDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
DB 666 LRSRPVVDLTVSMTFTP-----DALPMVFIGKM--DVNLCFEVDSSVVAEPGLREM 715
QY 654 QIQSVVTVDLALDSRPHRAVFNETKSTRTQVGLTQTC-----696
DB 716 FLNFTVDV-----TKQRQLQCEDSSGCSCLKRWGSGFLCBHFWLI 760
QY 697 ETLKQLPNCIEDPSPVLRNLSLVGTPLSAFNGLR----PVLAEQAQLRFTALF--P 750
DB 761 STEEL-----CEDCFSNITIKTYE-----FOTSGRRDYNPTLI--DHYKEPSAIFOLP 809
QY 751 FEKNGNDNICDDLSITFSFMSLDCLVVGGPRFNTVTVVNDGEDSYRTQVTFFFLD 810
DB 810 YEKDCKNVFCIAEIQLTIN-ISQBELVVGTVKEVTMTNISLTNSGDSYMTNMAIYNPRN 868
QY 811 LSYRKVSTLQONRSORSHRLACESASSTEVSALKSTCSINHPIPPENSEVTFNITDV 870
DB 869 LQFKKI-----QKVPSPDVQCDPKPV----ASVLVNMCKIGHPII-KRSSVNVSVTWOL 918
QY 871 DSKASLGNKLLKANVTSENNMPTNKTEFQLELPVKVAVYVMTSHGVSTKYLNFATASE 930

RESULT 8

ITAE_HUMAN STANDARD; PRT; 1179 AA.
ID ITAE_HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).
DE antigen (CD103 antigen) (Integrin alpha-IEL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RP TISSUE=Leukemia, and Lymphocytes;
RC MEDLINE=94164962; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;
RA "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
[2]
RN REVISIONS TO 88-114.
RP Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE OF 53-1179 FROM N.A.
RP TISSUE=Fetal kidney;
RC MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Aniketer Y., Dietrich N.L., Maduro V.V., McDowell G., Shetlerauk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;
RA "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";
RT Genome Res. 10:165-173(2000).
[4]
RN MUTAGENESIS OF ASP-109 AND PHE-316.
RX MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;
RA "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";
RT integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
 CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
 CC EPITHELIAL CELLS.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
 CC -!- SIMILARITY: NAME=PROW; NOTE=CD guide CD103 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L25851; AAB59359.2; -
 CC EMBL; AF168787; AAF431107.1; -
 CC PIR; A53213; A53213.
 CC HSSP; P11215; IABX.
 CC Genew; 604682; -
 CC MIM; 604682; -
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR002035; VWF_A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00357; Integrin_A; 1.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00191; Int_alpha; 3.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC PROSITE; PS00234; VWA; 1.
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Polymorphism; Magnesium; Calcium.
 FT SIGNAL 1 18
 FT CHAIN 19 1179
 FT CHAIN 19 1179
 FT CHAIN 19 1177
 FT CHAIN 179 1177
 FT CHAIN 19 1124
 FT TRANSMEM 1125 1147
 FT DOMAIN 1148 1179
 FT DOMAIN 181 198
 FT REPEAT ? ?
 FT REPEAT ? ?
 FT REPEAT ? ?
 FT DOMAIN 145 199
 FT DOMAIN 200 391
 FT REPEAT 401 456
 FT REPEAT 457 506
 FT REPEAT 510 571
 FT REPEAT 573 638
 FT REPEAT 641 693
 FT CA_BIND 522 530
 FT CA_BIND 586 594
 FT CA_BIND 654 662
 FT SITE 1150 1154
 FT DISULFID 70 79
 FT DISULFID 126 159
 FT DISULFID 706 762
 FT DISULFID 823 829
 FT DISULFID 893 907
 FT DISULFID 1008 1033
 FT DISULFID 1041 1057
 FT CARBOHYD 49 49
 FT CARBOHYD 271 271
 FT CARBOHYD 321 321
 FT CARBOHYD 444 444
 FT CARBOHYD 726 726

FT CARBOHYD 782 782
 FT CARBOHYD 857 857
 FT CARBOHYD 934 934
 FT CARBOHYD 954 954
 FT CARBOHYD 1065 1065
 FT CARBOHYD 1096 1096
 FT VARIANT 360 360
 FT VARIANT 1041 1041
 FT MUTAGEN 208 208
 FT MUTAGEN 316 316
 FT CONFLICT 477 477
 FT CONFLICT 482 482
 FT CONFLICT 950 950
 FT CONFLICT 1019 1019
 FT SEQUENCE 1179 AA; 130088 MW; E558902EDFD95E1 CRC64;
 SQ
 Query Match 19.2%; Score 1128; DB 1; Length 1179;
 Best Local Similarity 28.7%; Pred. No. 6.3e-69;
 Matches 338; Conservative 213; Mismatches 454; Indels 172; Gaps 39;
 QY 45 GSLYQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPQLLACQPTVHQ 95
 DB 65 GPLHRCSLVQDEILCHPVEHPVPKGRHGRVTVVRSHGVLCI-----QVLVRRP--HS 117
 QY 96 TCSENTYVKGCLFLGSLNRLOPQ-----SDIAFLIDGSGSIIPHDFRMKELVSTIMEQL-- 119
 DB 118 LSSELT---GTCSLGLPDLRPOAQANFFDLENLLDPDARVDTGDCYSNKEGGEDDVNTA 174
 QY 120 KPPEALRGCPQED-----SDIAFLIDGSGSIIPHDFRMKELVSTIMEQL-- 164
 DB 175 RQRALKEKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEED 234
 QY 165 KSKTLPFLMQLMSEEFRIHFTKEFQNNPNPSLKPITQLGRTHATGLRKVYRELFN 224
 DB 235 KCFECNFAVQGGVIQTEFLDRSDQVMASLARYQNTQVGSVTKTASAMQHVLDISFT 294
 QY 225 ITNGARKNAFKILFLLTDGKFGDPLGYEDVPELDREGVIRYVLGFGDAPRSEKROEL 284
 DB 295 SSHGSRKASKVMVLTGGIFEDPLNLTVINSPMQGVKVERFAIGVGEFKAARTAREL 354
 QY 285 NTVASKPPRDHVFQANNFEALKTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAIT 344
 DB 355 NLIASDPDETHAFKVTNYMALDGLLSKLRNIIISMEGT---VGDALHYQLAQIGFSAQIL 411
 QY 345 SNGP--LLSTVGSYDWAGGVLY--TSKEKSTFINMTRVSDMDNA-----YLGVAIAIILRN 398
 DB 412 DERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTAATAAADAEEAAQYSLGYAVAVLHKT 471
 QY 399 RVQSLVLGAPRYOHIGLVAMFR--QNTGMWESNANY--KGTQIGAYFGASICSVDVDSNGST 456
 DB 472 CSLSVAGAPQYKHG--AVFELQKEGREALFPLVLEGEQMGVSFGSELCPVDDIMDGGST 529
 QY 457 DLVLIGAPHYEQTRGGQVSVCPPLRGQARMQCDVAVLYGEOQPGWGFAGAAITVLGDVN 516
 DB 530 DFLVAAPFYHVHGEGRVYVRLSE--QDGSFSLARILSGHGFNARFGFAMAMGDLS 588
 QY 517 GDLTDTVAIGAP-----GEEDNR--GAVYLFHGTSGSGISPSHSQRIAGSKUSPRLOYFQ 570
 DB 589 QDKLTDVAIGAPLEGFGADGASFGSVIYNG--HWDGLSASPSQIRASTAVAPGLQYFGM 647
 QY 571 SILSGGQDLTMDGLVDLTVGAGOHVLLLRSPQVLRVKATMEFNPREVARNVFECNDQVVK 630
 DB 648 SMAGGFDISGGGLADITVGTGQAVFRSRPVRVLKVSNAFTPSALP-----IGF 697
 QY 631 KEAGEVVRVCLHVQKSTRDLREGQIQSVVTVYDLALDSGRPHSRVAFNETKNSRRTQVTL 690
 DB 698 NGVNVRLCFEI--SSVTTASESGLEALLNFTLDVDVGKQRRRLQCSVRSCLGCLREWS 756
 QY 691 GLTQTCETILKQLPN-----CIEDPVSPIVLRNLSLVTGTPLSAFGNLRLPVLAEQAQLFT 746
 DB 757 SGSQLCEDL--LLMPTGEGELCEBDCFSNASVKVSYQL--QTPEGQTDHPQILDTRYEPFAI 814

QY 747 ALFFPEKNCNDNICODLSTFMSFLDCLVWGPGPRENVTVTVRDNGEDSYRTQVTF 806
Db 815 FOLPYEKACKNKLFCVAELQLA-TVSQOELVGLTKELTLNLTNSGDSYMTSMALN 873
QY 807 PPLDLSYKVKSTLQNRQSRQSWRLACESASSTEVSGALKSTSCSINHIPPENSEVTFNI 866
Db 874 YPRNLQ-----LKRMRQPPSPNIQCDDPPQV-ASVLIMNCRIGHPLV-KRSSAHVSV 923
QY 867 TFDVDSKASLGNKLLKANVTSENN---MPTNKTFO---LPLPKYAVYVMTVSHGV 919
Db 924 VQUSEAFNRTADITVTNSMERRSLANETHLQRFHGFVAVLKSPSIWYNTQGL 983
QY 920 S--TKYLNFTASENTRVMOHQYQVNSLQSRSLPSLVFLVPLVRLNQTIVWDRPQVTFSE 977
Db 984 SHHKEFLPHVGEN---LFGAEYQ-----LQICVPTKLGRLQVAAVKLTQTQ 1028
QY 978 NLSSTCHTKERLPSHDSFLAELRKAPVNCVQRCIORQDIPFGQBEFNATLKNLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQHVVEWHVSVSCVIA-----SDKENVTVAABIS 1073
QY 1038 FDMYIKTSHNHLIVST-----AELFENDSVFTLLPGQAFVRSQETETKVPPEVNPPL 1091
Db 1074 WD-----HSELKDVTELOLGEISFNKSLYEGLNAENH--RTKITVFLKDEKYHSL 1125
QY 1092 PLIVGSSVGGLLALLITALYALYKLGFFKQYKQDMMSSE 1128
Db 1126 PLIVGSSVGGLLVILVILFKCGFFKQYKQDMMSSE 1162
RESULT 9
ID ITAI HUMAN STANDARD; PRT; 1151 AA.
AC P56139;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAL.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit".
RL J. Biol. Chem. 268:2989-2996 (1993).
CC - FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
E-R IN COLLAGEN.
CC - SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC - SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC - SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC - SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC - DATABASE: NAMB=PROW; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm".
CC PIR; A45226; A45226.
DR PDB; 1QC5; 17-MAY-00.
DR Genew; HGNC:6134; ITGAL.
DR MIM; 192968; -
DR GO; GO:0008305; C.integrin complex; TAS.
DR GO; GO:0004895; F.cell adhesion receptor activity; NAS.
DR GO; GO:0005518; F.collagen binding activity; TAS.
DR GO; GO:0007160; P.cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.

DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Repeat; Calcium; Magnesium; 3D-structure.
FT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1114 1136 POTENTIAL.
FT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GFFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;

Query Match 18.5%; Score 1085.5; DB 1; Length 1151;
Best Local Similarity 27.6%; Pred. No. 4.9e-66;
Matches 342; Conservative 209; Mismatches 488; Indels 201; Gaps 44;

QY 1 FNLDTENAMTFQENARG-FGQSVVQL---QGSRVVVGAPQEIIVANQSGSLYQCDYSTGS 56
Db 1 FNVVDVQNSMTFSGPVDPMFGYVQOYENEEKWVLIGSPLVGQPKNRTGVDYKCPVGRGE 60
QY 57 CEP-IRLQVPEA-----VNMSLGLSLAAATTPPQLLACGPTVHQTCSENTYVKG 106
Db 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGGFLACGFLAYRCGLHLYTTGI 119

Qy	107	CFIFGSLRQOPKPFBEALRGCPQEBSDIAFLIDGSGSIIPHDFRMKELVSTIMSQLKK	166
Db	120	CSDVSPFTQVQVNSIAP--VOECSTQ-LDIVILVDGNSIYPWD-----SVTAFNLDLLKR	171
Qy	167	-----SKTFLSLMOYSEEPRIHETFEFQNNPNPSRLIKPITOLLGR--THATGLRKVVR	220
Db	172	MDIGPKQTQVIGVOYGBNVTHENLNKYSISTEVLVAKKIVGGRQNTWALGTOTARK	231
Qy	221	ELFNITNGARKNAFKILFLITDGEKFGDPLGYEDVIPELDREGVIRYVLGFGDAFR---	276
Db	232	EATEARGARRGVKKVMVIVTDGESH-DNHRLKVKIQDCEDENIQBFSTAILGSYNRGNL	290
Qy	277	-SEKSRQELMNTVASKPRDRHVFOANPEALKTVQONLREKIFAEGTQTGSSSPFHEMS	335
Db	291	STEFVBEIKSIASEPTKEKHFNVSDBALVITVKTLGERIFALEATDAQSAASPHEMS	350
Qy	336	QEGFSAALTSNGPLLSITVGSYDRAGGVFLYSKE-----KSTP-INMTRVDSMDNDAYLG	389
Db	351	QTGFSAHYSQDWMLGAVGAYDNWGTVMVOKASOIIPRNTTTFNVESTKNEPL-ASYLG	409
Qy	390	YAAAILLRNRVQSL-VLGAPRYOHIGLVAFRQNTGMWESNANVKGTQICAYGASLCSV	448
Db	410	YTVNSATASGDVLIYAGORYNHTGOVIYRMEDGNIKILQTLSEQIGSYFSGILTTT	469
Qy	449	DVDSNGSTDVLIGAPHY-----YEOTR-GGQVSVCPLPRGQBARWQODA	492
Db	470	DIDKDSNTDILLGAPWYMGTEKEEQKVYVALNOTRFEYQWLSPEIKQTCSSRQHSN	529
Qy	493	VLGEQOQPNG-RFGAALTVLGDVNGDKLTDVAIGARGEDNRGNAYVLPHGTSGSGISPS	551
Db	530	CTTENKNEPCGARFGTAIAAVKDLNLDGFNDIVIGAPLEDDHGGAVYIYHG-SGKTIRKE	588
Qy	552	HSORLAGSKLSPLROYFGQSLSGQDILTMDGLVDLTVGAOCHVLLLRSQBVLKVAIMEF	611
Db	589	YAQRIPSGGDGKTLKFGQSIHEMDLNGDLTDTVIGGLGGAALFWSRDVAVVKTMWF	648
Qy	612	NPREVARNVPECNDQVVKGEAG--EVRVCLHVO-KSTRDLREGQIQSVVTVYDLALDSG	668
Db	649	EPNKVNIQKNCH---MEGKETVCINATVCFEVKLKSKEDTIYEADLQ---YRVTLDSL	701
Qy	669	RPHSRVFNET-----KNSTRQQTQVLGTQTCETLKLQPCNIEDPVSPIVLRNLSL	722
Db	702	QOISRSFSGTQERKVGORNTIVRKSEC-----TKHSFYMLDKGDFQDSVR---ITLDFNL	753
Qy	723	VGTPLSAFGNLRPVLAEDAQRLFTALPPFEKNGCNDNICDDLSITFSFMSLCLVVGVP	782
Db	754	T-DPENG-----PVLDDSLNSVHEYIIPFAKDCNKEKCLISLSLHVATTEKOLLIVRSQ	807
Qy	783	RE-FNVTVTVRNDGEDSYRTQVTFPPFDLSLYRKVSTLQNRSQSRWRLACESASSTVS	841
Db	808	NKDFNVSLTVKNTKDSAYNTRTIYHVSPLNVFSGIEAIQD-----SCESN-----	853
Qy	842	GALKSTSCSINHIFPENSEVTNIPFDVDSKASLGN-KULLKANTSENMMRPTNKTBF	900
Db	854	---HNITCKGVFPPLRGENMTFKLPQFNSTYLMENVTIYLSATSDSEBPPTLSDNVV	910
Qy	901	QLELPVKYAVYVMTVSHGVSTKYLVNFTASENTSRVMQHQVQSN-----LGORS----	949
Db	911	NISIPVKYEVLOFYS-SASEYHISIAANTVEVINSTEDIGNEINIFYLIRKSGSPM	969
Qy	950	--LPISLVF-----LVPVRLNQTIVDRPQVTFSENLSSTCHTKE-----R	988
Db	970	PELKLISIFPNMTSNGYPVLYPTGLSS-----SENANCRPHIFEDPFSINGKK	1018
Qy	989	LPSHSDPLAELKAPVNVCSIAVCORIQCDIPFGTQE-----EFNA	1030
Db	1019	MTTSTD---HLKRGITLDCNCTCKFATITCLNTSSDISQVNVSLILWKPTFIKSYFSSLNL	1075
Qy	1031	TLKGNLSFDWYKITSNNHLIIVSTAEILFNDSVFTLLPGOGAFVRSQTETKVEPFVBNP	1090
Db	1076	TIRGEL-----RSENASIVLSSN-----OKRELAIQISKOGLPGR	1111
Qy	1091	LPL--IVGSSVGGILLLIALITAAALKYKLGFFKROVKDMWSE	1128

```

Db          1112 VPLWVILLSAFAGLLMLLILALWKIGFKRPLKKKMK 11151
      :||| |:|: |:| ||||| |:| |:|:||||| | |:| :|
RESULT 10
ITAH_HUMAN
ID        ITAH_HUMAN    STANDARD;             PRT;   1189 AA.
AC        Q9UKX5; Q9UKQ1;
DT        16-OCT-2001 (Rel. 40, Created)
DT        16-OCT-2001 (Rel. 40, Last sequence update)
DT        28-FEB-2003 (Rel. 41, Last annotation update)
DE        Integrin alpha-II precursor.
GN         ITGAI1.
OS        Homo sapiens (Human).
OC        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX        NCBI_TaxID=9606;
RN         [1]
RC        SEQUENCE FROM N.A.
RX        TISSUE=Fetal heart, and Osteoblast;
RX        MEDLINE=99417678; PubMed=10486209;
RA        Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yacobi Z., Wang S.-X., Morris C.M., Krissansen G.W.;
RA        "Cloning, sequence analysis, and chromosomal localization of human integrin alphai1 subunit (ITGAI1).";
RT        Genomics 60:179-187(1999).
RL         [2]
RC        SEQUENCE FROM N.A.
RX        TISSUE=Fetal muscle, and Uterus;
RX        MEDLINE=99395147; PubMed=10464311;
RA        Velling T., Kusche-Gullberg M., Sejersen T., Gullberg I., "cDNA Cloning and Chromosomal Localization of Human alpha-1 Integrin. A collagen-binding, i domain-containing, beta-1 integrin alpha-chain present in mouse tissues.";
RT        J. Biol. Chem. 274:25735-25742(1999).
RL         [3]
RC        SEQUENCE OF 954-1188 FROM N.A.
RP        Tissue=Fibroblast;
RA        Andreu N., Estivill X., Escarceller M., Smey L.;
RL        Submitted (JUN-2000) to the EMBI/GenBank/DBJ database.
CC        -1- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT ASSOCIATES WITH BETA-1.
CC        -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC        -1- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING TO SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIAL MUSCLE CELLS (IN VITRO).
CC        -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DO WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC        -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY. SIMILARITY: Contains 1 VWFA domain.
CC        -1- SIMILARITY: Contains 7 FG-GAP repeats.
-----
CC        This SWISS-PROT entry is copyright. It is produced through the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no use by non-profit institutions as long as its contents modified and this statement is not removed. Usage by CC entities requires a license agreement (See http://www.ebi.ac.uk/send an email to licensee@sib-sib.ch).
-----
DR        ENBL; AF109681; AAF01258.1; -.
DR        ENBL; AF137378; AAD51919.2; -.
DR        ENBL; AL359064; CAB94392.1; -.
DR        HSHP; PL7303; IAOX.
DR        DR        Gnew; HGNC:6136; ITGAI1.
DR        DMN; 604789; -.
DR        GO; GO:0008305; C:integrin complex; TAS.

```



```
QY 995 FLAELRKAPVNCISIAVCORIQCDIPFGIOBEFNATLGNLSFDVY-----IKTSHNHL 1049
Db 1051 --EDLRAPQLNHSNDVVSINCIRLVP-NQEIFHLLGNL---WLRSLKALKYKSMKI 1104
QY 1050 LIVSTAEILLFNSVFTLLPQGAFAVSOTETKVERPEVEN-----PLPLIVGSSVGGLL 1103
Db 1105 MVNAALQRFH-SPF-----IFREBPSRQIVFEISKQEDWQVPIIIVGSTLGGLL 1155
QY 1104 LIALITALYKLGFFK--RQYKMMSEGGPPGAEP 1136
Db 1156 LLALLVLWLKLGFFRSARRRE-----PGLDP 1183

RESULT 11
ID ITA2_BOVIN STANDARD; PRT; 1170 AA.
AC P53710;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).
GN ITGA2;
OS Bos taurus (Bovine).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94193647; PubMed=7511592;
RA Kamata T., Puzon W., Takada Y.;
RT "Identification of putative ligand binding sites within I domain of
RT integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
RL J. Biol. Chem. 269:9659-9663(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ.
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL-outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L25886; AAB59255.1; -
DR FIR; I45914; I45914.
DR HSP; P17301; IA0X.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP_3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
```

```
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
FT NON_TER 1
FT SIGNAL <1 18
FT CHAIN 19 1170
FT DOMAIN 19 1121
FT TRANSMEM 1122 1143
FT DOMAIN 1144 1170
FT REPEAT 34 92
FT REPEAT ? ?
FT DOMAIN 177 367
FT REPEAT ? ?
FT REPEAT 423 475
FT REPEAT 477 538
FT REPEAT 540 599
FT REPEAT 604 656
FT CA_BIND 488 496
FT CA_BIND 552 560
FT CA_BIND 616 624
FT SITE 472 474
FT SITE 1146 1150
FT DISULFID 72 81
FT DISULFID 669 726
FT DISULFID 778 784
FT DISULFID 854 865
FT DISULFID 1008 1039
FT DISULFID 1044 1049
FT CARBOHYD 94 94
FT CARBOHYD 101 101
FT CARBOHYD 332 332
FT CARBOHYD 421 421
FT CARBOHYD 449 464
FT CARBOHYD 454 464
FT CARBOHYD 688 688
FT CARBOHYD 748 748
FT CARBOHYD 945 945
FT CARBOHYD 1063 1063
FT CARBOHYD 1070 1070
FT VARIANT 580 588
FT VARIANT 588 588
FT VARIANT 725 725
SQ SEQUENCE 1170 AA; 128929 MW; BECF1C5F2448PB1 CRC64;

Query Match 18.2%; Score 1071; DB 1; Length 1170;
Best Local Similarity 27.5%; Pred. No. 5e-65;
Matches 334; Conservative 218; Mismatches 495; Indels 168; Gaps 48;

QY 1 FNLDTENAMTQ-ENARFGSGSVQL---QGSRVVVGAPQEIIVAAQNGSLYQC--DYST 54
Db 19 YNVGLPKAKIFGSPSSEQFGYAVQQFINPKGNWLLVSGPWSGFPKRMGDVYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGLC 107
Db 79 TTCEKLNLOTSTMSNSNVTMTNMSLGLTLRNVTGCGFLTCGPLWAQCGSQYTTGVC 138
QY 108 FLFGSNLRQQQPKPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLK-- 165
Db 139 SDVSPDF-QLRTSPAPAVQTCP-SPIDVVVWCDESNIYPMB--AVKNFLKFXVQGLDIG 194
QY 166 KSKTLFSLMYSSEPRHIFTKFEQNNPNPSLIKPIQLL-----GRTHATGLRKVRE 221
Db 195 PTKTMGLIQIYANNPRVVFNLNTPFKSKD---EMIKATSQTFQYGGDLTNFKAIQYARD 251
QY 222 LPNITNGARKNAFKILFLTLTGCKPGDPLGVEDVPELDREGVIRY--VLGF--GDAFR 276
Db 252 AYSTAAGRGPGATKVMVVVTIDGESH-DGSKLKAVIDQCNKONILRFGTAVLYLRNALD 310
QY 277 SEKSROELNTVASKPPRDHVFQANNFEALKTQVQNLREKIFAIEGTQTGTGSSSFEHMSQ 336
Db 311 TKNLIKEIKATASIPTEHRHFNVSDADLLEKAGTIGQIFSIETGVQG-GDNFQWMSQ 369
QY 337 EGFSAAIT--SNGPLLSTVGYDWAAGGVFLYTSKEKSTFINMT--RVSDSMV-DAYLGYA 391
```

```
Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
NON_TER 1
SIGNAL <1 18
CHAIN 19 1170
DOMAIN 19 1121
TRANSMEM 1122 1143
DOMAIN 1144 1170
REPEAT 34 92
REPEAT ? ?
DOMAIN 177 367
REPEAT ? ?
REPEAT 423 475
REPEAT 477 538
REPEAT 540 599
REPEAT 604 656
CA_BIND 488 496
CA_BIND 552 560
CA_BIND 616 624
SITE 472 474
SITE 1146 1150
DISULFID 72 81
DISULFID 669 726
DISULFID 778 784
DISULFID 854 865
DISULFID 1008 1039
DISULFID 1044 1049
CARBOHYD 94 94
CARBOHYD 101 101
CARBOHYD 332 332
CARBOHYD 421 421
CARBOHYD 449 464
CARBOHYD 454 464
CARBOHYD 688 688
CARBOHYD 748 748
CARBOHYD 945 945
CARBOHYD 1063 1063
CARBOHYD 1070 1070
VARIANT 580 588
VARIANT 588 588
VARIANT 725 725
SEQUENCE 1170 AA; 128929 MW; BECF1C5F2448PB1 CRC64;

Query Match 18.2%; Score 1071; DB 1; Length 1170;
Best Local Similarity 27.5%; Pred. No. 5e-65;
Matches 334; Conservative 218; Mismatches 495; Indels 168; Gaps 48;

1 FNLDTENAMTQ-ENARFGSGSVQL---QGSRVVVGAPQEIIVAAQNGSLYQC--DYST 54
19 YNVGLPKAKIFGSPSSEQFGYAVQQFINPKGNWLLVSGPWSGFPKRMGDVYKCPVDLST 78
55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGLC 107
79 TTCEKLNLOTSTMSNSNVTMTNMSLGLTLRNVTGCGFLTCGPLWAQCGSQYTTGVC 138
108 FLFGSNLRQQQPKPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLK-- 165
139 SDVSPDF-QLRTSPAPAVQTCP-SPIDVVVWCDESNIYPMB--AVKNFLKFXVQGLDIG 194
166 KSKTLFSLMYSSEPRHIFTKFEQNNPNPSLIKPIQLL-----GRTHATGLRKVRE 221
195 PTKTMGLIQIYANNPRVVFNLNTPFKSKD---EMIKATSQTFQYGGDLTNFKAIQYARD 251
222 LPNITNGARKNAFKILFLTLTGCKPGDPLGVEDVPELDREGVIRY--VLGF--GDAFR 276
252 AYSTAAGRGPGATKVMVVVTIDGESH-DGSKLKAVIDQCNKONILRFGTAVLYLRNALD 310
277 SEKSROELNTVASKPPRDHVFQANNFEALKTQVQNLREKIFAIEGTQTGTGSSSFEHMSQ 336
311 TKNLIKEIKATASIPTEHRHFNVSDADLLEKAGTIGQIFSIETGVQG-GDNFQWMSQ 369
337 EGFSAAIT--SNGPLLSTVGYDWAAGGVFLYTSKEKSTFINMT--RVSDSMV-DAYLGYA 391
```

```
Db 370 VGFAEYSPONNMLGAVGDMGTGVVQKTPHGHLLIFSKOAFEQILQDRNHSSYLGS 429
Qy 392 RAAILLRNVQSLVLCAPRYQHTGLVAMPRQNTGMWESNAV-----KGTOIGAYFGASL 445
Db 430 VASISTGNSVHFVAGAPRANVTGQIVLVSX-----ENGNTVIOQRGDQIGSVFGSVL 484
Qy 446 CSVDVDSNGSTDVLIGAPHYYEOTR--GGQVSVCPPLPRGQRARQCDAVLYGQGPWG 503
Db 485 CAVDVNKDITDVLVGVAPMYNDLKKEGRVYLTITKG-ILNWH--QFLEGPNGLENA 541
Qy 504 REGAALTVLGDVNGDKLTDVAIGAPGEDNRCNAVLFPHGTSGSGISPSHSQRIAGS--KL 561
Db 542 REGSAIALSDINMDGDFNDVIVGSPLENQNSGAVIYNGHEGM-ILRYSOKILGSDRAF 600
Qy 562 SPRLQYFGOSLGGQDLTMDGLVTLVGAQGHVLLRQPVLRVKAIMEFNPREVARNVF 621
Db 601 SHLQYFGESLDYGLDNGSDITDVSAGFGVQVQLMSQSDADVSVDASFTPKKI--TLL 658
Qy 622 ECNDQVVKGEAGEVRVLHVQKSTRDLRREGQIOSVVTYDALP-----SGRPHRAVFN 677
Db 659 NKNABE-----KLKLCF---SAKFRPTNQNNQVAIYNTIDEDQFSRVRISGLFK 707
Qy 678 ETKNSTRTQVLGLTQICE--TLKQLPNCIEDPVPVILRLNFSL--VGTPLSAFGLN 733
Db 708 ENNERCLQKTMIVSQACRSEYIIHQEPS-----DIISPLNLCMNISLENPGT----- 756
Qy 734 RPLVAEDAORLFTALFPPEKNGCNDNICODLSITF----SPMSLDCLVGVGPRFNVTV 789
Db 757 NFALAEYSTVKVFPFPHKCGDGVCLSDVLNVQLPATQOPQFFVSNQNKLTFSV 816
Qy 790 TVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWLACESAST-EVSGALKSTS 848
Db 817 QLNKKESAYNTEIVVDFSENLF-----ASWMPVDGTEVTQCIASSQKSVT 864
Qy 849 CSINHIPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEFQLELPVKY 908
Db 865 CNVGPALKSKQVTFITNFDNLQ-NLQNASISFRALSESQEEENMADNSVNLKSLLY 923
Qy 909 AVYVMTVSHGVTKYLNFTASENTSRVMOHOYQVSNLQOR-----SLPISLVLFLV 958
Db 924 DAEIHT-RSTNINFEVSLDGNVASV-HSFE--DIGPKFTFSIKVTGSPVPSWA--- 976
Qy 959 PVRLNQTVIMDRPQVTFSEN-----LSTCHTKS-----RLPSHSDFLAE- 998
Db 977 -----SVIIHIPQVTKDNPLMYLTGVHTDQAGDISCEAINPLKIGTSSSVFSKSEN 1030
Qy 999 LRKAPVNCIAVCQBIQCDIPFGIQEFPFNATLKNLSFDWYIKTSHNHLIVSTAET- 1057
Db 1031 FRHIKELNCRATASCSNIMCWLRLQVKGEYFLNVSTRIWNGTFAASTFOTVOLTAAAEID 1090
Qy 1058 LFNDVSFTL-----LPQGFVRSQETETKVEPFE-VPNPLPLVIGSSVGGLLALLALITA 1110
Db 1091 TYNPQIYVIEENTVTP-----LTIKPHKEVETPTGVIVGSVIAGIALLALLALVA 1140
Qy 1111 ALYKLGFFKRYQKDM 1125
Db 1141 ILWKLGFKKRYEKM 1155

RESULT 12
ID ITA2 MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

FT CA_BIND 560 568 POTENTIAL.
 FT CA_BIND 624 632 POTENTIAL.
 FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1154 1158 GFKR MOTIF.
 FT DISULFID 80 89 BY SIMILARITY.
 FT DISULFID 677 734 BY SIMILARITY.
 FT DISULFID 786 792 BY SIMILARITY.
 FT DISULFID 862 873 BY SIMILARITY.
 FT DISULFID 1016 1047 BY SIMILARITY.
 FT DISULFID 1052 1057 BY SIMILARITY.
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 457 472 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 496 696 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1178 AA; 128926 MW; IF194B9C0240F465 CRC64;
 Query Match 18.0%; Score 1057; DB 1; Length 1178;
 Best Local Similarity 27.8%; Pred. No. 4.5e-64;
 Matches 341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;

QY 1 FNLDTENMTFQ--ENARFGQSVVQL---QGSVVVGAPOEIVAAQROGSLYQC--DYST 54
 Db 27 YNVLPGAKIFSGPSSQFGVYQQLTNPQGNMLLVGSPWSPFENRMDGVYKCPVDLPT 86
 QY 55 GSCEPIRLQ-----VPVEAVMSGLSLAATSPPOLACGTFVHQTSENYYKGLC 107
 Db 87 ATCEKUNLQNASISNVTETKTNMSGLTLTRNPGTGGFLTCPLMAHOCGNOYYATGIC 146
 QY 108 FLFGSNLRQOPQ---KPEALRGCPQSDIAFLIDGSGSIIIPHDFFRMKELVSTIMEQL 164
 Db 147 ----SDVSPDFQELTSPVAVQCPSL--VDVVVVCDESNIYP--WEAVKNFLVKPVTGL 199
 QY 165 K--KSNTLFSLMQYSEBRFHFTFKFQNNPNRSLIKPIITQLLG-RTHPTATGLRKVVRE 221
 Db 200 DIGPKTKTQVALIOYANPRIFINLNDPTEKMDVQATSETROHGGDLTNTFRAIEFARDY 259
 QY 222 LFNITGARGNAKILFLLLTDGKFGDPLGYEDVIELDREGVIRY---VLGF--GDAFR 276
 Db 260 AYSQTSGRGRGATKVMVVVDGESH--DGSKLKTVIOQCNDEDLRFGIAGVLGYLNRLALD 318
 QY 277 SEKSQRLNVTASKPPRDHVFQANNFEALTKVQNLREKIFAEJGTGTGSSSSFEHMSQ 336
 Db 319 TKNLKIKKAIATPTETERYFENVADAEALKEKAGTLGEQIFSIETVQG--GDNFQEMAQ 377
 QY 337 EGFSA--AITSNGPLLSVGSYDWAGGVFLYTSKEKSTFINMT--RYVDSMN--DAYLGVA 391
 Db 378 VGFSAADYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIFPKQAFDQVLQDRNHSFLGYS 437
 QY 392 AAILNRVOSLVGAPRYOHIGLVANFRONTGWESNANV----KGTQIGAVFGASLCS 447
 Db 438 VAAISTBDGVHFVAGAPRANYTGQVLYSVNK---QGNVTVIOQSHRGDQIGSYFGSVLCS 494
 QY 448 VDSDNSGTDVLILGAPHYEQTR--GGQVSVCLPREGQARQCCAVLYGEGQQPWGRF 505
 Db 495 VDVKOTITDVLVGATYHNDLKEEGKYLFTITKGIUNHQ---FLEGPGTGNARF 551
 QY 506 GAALTVLGDVNGDKLTVAIGAPCEEDNRGAVLYFHTGTSKGISPSHQSRIAGSKLSR- 564
 Db 552 GSAIALSDINMDGFNDIVGSPVENENSGAVIYNGHQT--IRTKYSQKILSGNGAFRR 610
 QY 565 -LQYFGSLSGGQDLTWGLVDLTVGAGQHVLRLRSQPVLRVKAIMEFNPVARNVFE 623
 Db 611 HLOFFGRSLDGYGLNDGSDITDVSIGALGQVIOQMSQSIADVAIEALFTP----- 660
 QY 624 NDQVVKGEAGEVRVCLHVQKSTDRLRREGOIQSVVTDLALD----SGRPHSRVFN 679
 Db 661 -DKITLUNKAKITLKLCFRAERFPAGQNNQV--AILFNMTLADAGHSSRVTSRGVFR 717

QY 680 KNSTRQTOVLGLTQTCET--LKLQLPNCIEDPVPDIVLRNPLSVLGTPLSAFGNRPVL 737
 Db 718 SERFLQKNVNVNEVQKSEHHSIQKPS---DVVNPDLRLVDISLENPGTS-----PAL 768
 QY 738 AEDAQLFTALPFPEKNCNNDNICODDLSI-----TFSPMSLDCLVVGCPREFNVT 788
 Db 769 EAYSETVKVFSIPFYKECGSDGICISDLILDVQQQLPAIQOTQSF-----IVSNQKRLTFS 823
 QY 789 VTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACBSASST-EVSGALKST 847
 Db 824 VILKNRGESAYNTVVLAEFSENLF-----ASFSPVDGTEVTCVSGSQKSV 871
 QY 848 SCGINHIPENSEVETNITFDVDSKASLGNKLLKANVTSENMPRTNKT--FOLELP 905
 Db 872 TCDVGYPALKSEQQVTFETINFENLQ--NLQQAIAINFOAFSESQ--ETNKADNSVSLTIP 928
 QY 906 VKYAVVMVVTSHGVSTKYLNFETASNTSRVMQHQVQVSNLQOR-----SLPISLV 955
 Db 929 LIYDAELHUT-RSTWINFYEISDENAPSVIK---SVEDIGPKFIFSLKVTAGSADPVMA 984
 QY 956 FLV-----PVLNQTVIWRPQVTF-SENLS 980
 Db 985 LVTIHIPQVTKKKNPLLYLTGIQTQAGDISCTAEINPLKPHTA-----PSVSFKNENR 1040
 QY 981 STCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIPFGIOEEFNATLKNLSFDW 1040
 Db 1041 ---HTKE-----LDCRTTSCNITCMLKDLHMAEYFINVTTRVWNR 1080
 QY 1041 YIKTSHNHLIVSTAEILLFNDSPVFTLLPGQAFVRSQTKVETKPEPPEVNPPLPLVGVSSVG 1100
 Db 1081 FRASTPQTQVLTAAAEIDTHNPQLFVIBENAVTIPLIMMKTEKAEVPT--GVIIGSIIA 1138
 QY 1101 GLLLALITAALYKLGFFKRYQKDM 1125
 Db 1139 GILLALLMTAGLWKLGFKRYQKDM 1163

RESULT 13
 ID ITA2 HUMAN STANDARD; PRT; 1181 AA.
 AC P17301;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
 GN ITGA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
 RC TISSUE=Endothelial cells;
 RX MEDLINE=89308879; PubMed=2545729;
 RA Takada Y., Hemler M.E.;
 RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit (platelet GPIa): homology to other integrins and the presence of a possible collagen-binding domain.";
 RL J. Cell Biol. 109:397-407(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
 RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
 RX MEDLINE=98019223; PubMed=9353312;
 RA Emsley J., King S.L., Bergelson J.M., Liddington R.C.;
 RT "Crystal structure of the I domain from integrin alpha2betal.";
 RL J. Biol. Chem. 272:28512-28517(1997).
 RN [4]
 RP VARIANT HPA-5 (BR).

RA MEDLINE=94043762; PubMed=7901236;
 RA Santoso S., Kalb R., Walka M., Kiefel V., Mueller-Eckhardt C.,
 RA Newman P.J.;
 RT "The human platelet allantoicins Br(a) and Brb are associated with a
 RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
 RT alpha 2).";
 RL J. Clin. Invest. 92:2427-2432(1993).
 RN [5]
 RP VARIANT GLU-534.
 RX MEDLINE=20206009; PubMed=10744142;
 RA Kroll H., Gardemann A., Fechter A., Haberbosch W., Santoso S.;
 RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
 RT gene polymorphism on coronary artery disease and acute myocardial
 RT infarction";
 RL Thromb. Haemost. 83:392-396(2000).
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC EXTRACELLULAR MATRIX.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
 CC associates with beta-1. Interacts with HPSS.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- ALLOMORPHISM: POSITION 534 IS ASSOCIATED WITH PLATELET-SPECIFIC
 CC POLYMORPHISM HPA-5 (BR). HPA-5A/BR(A) HAS LYS-534 AND HPA-5B/BR(B)
 CC HAS GLU-534. HPA-5B IS INVOLVED IN NEONATAL ALLOIMMUNE
 CC THROMBOCYTOPENIA (NAIT OR NATP). THE K534E POLYMORPHISM MAY PLAY A
 CC ROLE IN CORONARY ARTERY DISEASE (CAD).
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49b.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X17033; CAA34894.1;
 DR EMBL; AF512556; AAM34795.1;
 DR PIR; A33998; A33998.
 DR PDB; 1A0X; 25-NOV-98.
 DR PDB; 1DZ1; 02-AUG-01.
 DR Genew; HGNC:6137; ITGA2.
 DR MIM; 192974;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0005518; F:collagen binding activity; TAS.
 DR GO; GO:0007596; P:blood coagulation; TAS.
 DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWFA.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRIN.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;

3D-structure. 1 29
 FT SIGNAL 30 1181 INTEGRIN ALPHA-2.
 FT CHAIN 30 1132 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1133 1154 POTENTIAL.
 FT TRANSMEM 1155 1181 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1155 1161 INTERACTION WITH HPSS.
 FT DOMAIN 1155 1161
 FT REPEAT 45 103 FG-GAP 1.
 FT REPEAT 45 103 FG-GAP 2.
 FT REPEAT 188 378 VWFA.
 FT DOMAIN 378 433 FG-GAP 3.
 FT REPEAT 434 486 FG-GAP 4.
 FT REPEAT 488 549 FG-GAP 5.
 FT REPEAT 551 610 FG-GAP 6.
 FT REPEAT 615 667 FG-GAP 7.
 FT CA_BIND 499 507 POTENTIAL.
 FT CA_BIND 563 571 POTENTIAL.
 FT CA_BIND 627 635 POTENTIAL.
 FT SITE 1157 1161 GFFKR MOTIF.
 FT DISULFID 83 92 BY SIMILARITY.
 FT DISULFID 680 737 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT DISULFID 865 876 BY SIMILARITY.
 FT DISULFID 1019 1050 BY SIMILARITY.
 FT DISULFID 1055 1060 BY SIMILARITY.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 699 699 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 534 K -> E (IN ALLOANTIGEN HPA-5B;
 FT dbSNP:1801106).
 FT /FTid=VAR_003977.
 FT TURN 170 171
 FT STRAND 173 180
 FT TURN 183 184
 FT HELIX 188 199
 FT TURN 200 201
 FT TURN 204 204
 FT STRAND 206 207
 FT TURN 206 207
 FT STRAND 209 216
 FT STRAND 220 224
 FT TURN 226 228
 FT HELIX 232 240
 FT TURN 241 241
 FT HELIX 252 262
 FT TURN 263 264
 FT HELIX 266 268
 FT TURN 269 269
 FT STRAND 275 282
 FT HELIX 289 291
 FT HELIX 292 301
 FT TURN 302 303
 FT STRAND 304 311
 FT HELIX 313 317
 FT TURN 318 319
 FT HELIX 323 330
 FT TURN 331 332
 FT HELIX 337 340
 FT STRAND 341 344
 FT HELIX 347 353
 FT HELIX 354 362
 FT TURN 363 363
 SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match 17.9%; Score 1054; DB 1; Length 1181;
 Best Local Similarity 26.8%; Pred. No. 7.3e-64;
 Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;

ITAG_HUMAN
 ID ITAG_HUMAN STANDARD; PRT; 1167 AA.
 AC O75578; Q9UHZ8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-10 precursor.
 GN ITGA10.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Articular chondrocytes;
 RX MEDLINE=98352078; PubMed=9685391;
 RA Camper L., Hellman U., Lundgren-Akerlund E.;
 RT "Isolation, cloning, and sequence analysis of the integrin subunit
 alpha10, a beta1-associated collagen binding integrin expressed on
 chondrocytes.";
 RL J. Biol. Chem. 273:20383-20389(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=endothelial cells, and heart;
 RX MEDLINE=20169197; PubMed=10702680;
 RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
 Wang S.-X., Langley R., Kriesaneen G.W.;
 RT "The integrin alpha10 subunit: expression pattern, partial gene
 structure, and chromosomal localization.";
 RL Cytogenet. Cell Genet. 87:238-244(1999).
 CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
 CC ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSIONS IN
 CC MUSCLE AND HEART. FOUND IN ARTICULAR CARTILAGE.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA repeat.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF074015; AAC31952.1; -;
 CC EMBL; AF112345; AAF21944.1; -;
 CC EMBL; AF172723; AAF61638.1; -;
 CC HSP; P17301; IAOX.
 CC Genew; HGNC:6135; ITGA10.
 CC MIM; 604042; -;
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC GO; GO:0005518; F:collagen binding activity; TAS.
 CC GO; GO:0007160; P:cell-matrix adhesion; TAS.
 CC InterPro; IPR000413; Integrin_alpha.
 CC InterPro; IPR020335; VWF_A.
 CC Pfam; PF01839; FG-GAP; 3.
 CC Pfam; PF00092; vwa; 1.
 CC PRINTS; PR01185; INTEGRINA.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00191; Int_alpha; 4.
 CC SMART; SM00327; VWA; 1.
 CC PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 CC PROSITE; PS0234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium.
 FT SIGNAL 1 22 POTENTIAL.

CHAIN 23 1167
 FT DOMAIN 23 1122 INTEGRIN ALPHA-10.
 FT TRANSMEM 1123 1145 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 38 97 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT DOMAIN 167 350 VWFA.
 FT REPEAT 365 427 FG-GAP 3.
 FT REPEAT 428 482 FG-GAP 4.
 FT REPEAT 483 545 FG-GAP 5.
 FT REPEAT 546 605 FG-GAP 6.
 FT REPEAT 608 660 FG-GAP 7.
 FT DOMAIN 1134 1140 POLY-LIU.
 FT CA_BIND 494 502 POTENTIAL.
 FT CA_BIND 558 566 POTENTIAL.
 FT CA_BIND 620 628 POTENTIAL.
 FT DISULFID 76 86 BY SIMILARITY.
 FT DISULFID 666 675 BY SIMILARITY.
 FT DISULFID 681 736 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT CARBOHYD 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 844 844 I -> L (IN REF. 2).
 FT CONFLICT 909 909 G -> V (IN REF. 2).
 FT CONFLICT 926 926 E -> D (IN REF. 2).
 SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;
 Query Match 17.7%; Score 1042.5; DB 1; Length 1167;
 Best Local Similarity 28.5%; Pred. No. 4.4e-63;
 Matches 348; Conservative 201; Mismatches 495; Indels 177; Gaps 44;
 QY 1 FNLDTENAMTPOENARG-FGQSVVQLQGR-----VVGAPQEIIVANQROSLVOC----- 50
 DB 23 FNLDHHPRLFGPPPEAFSGYSLVQVGGQGRWMLVGAPWDGPGSDRRGDVVRCPVGGAH 82
 QY 51 -----DYSTG-SCEPIRLOVPVEAVNMVLSLAATTSPPQLACGTVVQTCS 99
 DB 83 NAPCAKHLGDYQLGNSSHP-----AVNMHLGMSLETDGCGGFMACPLMWRACCS 134
 QY 100 NTYVKGCLFLFGSNLRQPKPEALRGCPQSDSDIAFLIDGSGSIIPHD-----FRMK 154
 DB 135 SVFSSGICARVDASFQPGSLAPTQR-CPTY-MDVIVLDGNSIYPHSEVQTFRLRLV 192
 QY 155 ELVSTIMEQLKSKTFLSLMQVSEBFRIFHTPKFQNNPNRSLKIPITQLLGR-THTAT 213
 DB 193 GKLFIDPEIQ-----VGLVQYGESPVHESLGDFTKEEVVRAAKNLSRREGRETAK 247
 QY 214 GLRKVVRELFNITNGARKNAFKILLLTGEKF-GDPLGYEDVIBELDEGVIRY-VLGF 271
 DB 248 AIMVACTGFGSQSHGRPEARLLVVVDGSHDGEEL--PAALKACEAGRVTRYGVIAVL 305
 QY 272 GDAPRSEKS-----ROELNVTASKPRDRHVQFQANNFEALKNQVQNRKIFAJEGTQTS 327
 DB 306 GHYLRQRDPSPFLREIRTIASDPERFFNVTDDEALTDIVDALGDRIFEGGSHAENE 365
 QY 328 SSFEHEMSQEGFSAAITNGPLLSLVGVSDMAGGVFLYTSKEKSTFINMTRVDS-----D 382
 DB 366 SSFGLMSQIGFSTHRLKDGILFGMVAGYDMGGSVLWLEGGHRLFPFRMALEDEFPALQ 425
 QY 383 MNDAYLGYA-AAIILNRVQSLVLCAPRYOHLGLVAMER-QNTGHWESNANVKGQIGAY 440
 DB 426 NHAAYLYGVSVMLLRGRRLLFLSGAPRFRHKGKVIATFQKKDGAVRVAQSLQEQIGSY 485
 QY 441 FGASLCSVDVDSNGSTDLVLICAPHY--EQTGRGQVSVCPPLRQORARWQCDVLYGQ 498

Db 486 FGSELCPDTRDGTDDVLLVAAPMFLGPONKETGRVYVYLV--GQOULLTQGTQLEP 543
Qy 499 QPWRGRCFAALTIVLGDVNGDKLTVAIGAPCEENRGAVYLFHGTSGSGISPSHSORAG 558
Db 544 PQD-ARFGFANGALPDLNQGDFADVAVGAPLEDGHQGALYLYHGTQ-SGVRPHPAQRIAA 601
Qy 559 SKLSPRLQYFGOSLSGGDLTMDGLVDLTGGAQGHVLLRSOPVLRVKAI MEENPREVAR 618
Db 602 ASMPHALSYFGSRVDGRDLDDGDDLDVAVGAQGAAILLSRPVHLTPSLEVTPOAISV 661
Qy 619 NVFECNDQVVKKEAG--EVRVCLHVQKSTRDRREGQIQSVVYDYDALDSGRPHSRAVF 676
Db 662 VORDCRR---RGQEAUCLTAALCEQVTSRTFCRWDH---QFYMRFTASLDEWTAGARAAF 715
Qy 677 NET--KNSTRQTVLGTQTCETLKLQPNCEIDPVSPIVLRNLSVLTGTPLSAFGNLR 734
Db 716 DSGQRLSPRLRLSVG-NVTCEQLHFHVLD-TSDYLRPVALTVTFALDNTTKFG----- 768
Qy 735 PVLAEDAQRLFTALPPFEKNCNDNI CODDLSITFSPMSLDC-----LVVGGPREFN 786
Db 769 PVLNEGSPSTQKLVFPFSGDKGPNCEVTDVLQ---VNMDIRSRKAPFVVRGRRKVL 825
Qy 787 VTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEVSGALKS 846
Db 826 VSTTLNKRKNAYNTSLIIFSRNL--HLASLTPQR-ESPIKVECAAPSA-----HA 874
Qy 847 TSCSINHPIFPENSEVTNITFDVDSKASLG---NKLL-----LKAVTSENMPRTNKT 898
Db 875 RLCSVGHVPVOTGAKVTELEFEFSCSLLSQVFGKLTASSDSLSERNGTQENTAQT--- 931
Qy 899 EFQLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVNHQV-----QVS 943
Db 932 -----SAYIOYEH-----LFFSSESTLHRYEYHPYGTLPVGPGEFKTTLVQ 975
Qy 944 NLG---QRSPLISLVLP-----VRLNQTVIMDRPQVTFSENLSSTCHTKERLPS 991
Db 976 NLGCYVWSGLIISA--LLPAVAHGGNYFLSLQVI-----TNNASCIVQNLTEPPG 1024
Qy 992 HSDFLAELRKAPVNCSTAVCQRIQCDIPFFGIQEEFNATLKGNLSFDWYIKTSHNLLI 1051
Db 1025 PPVHPEELQHTNRLNGSNTQCVVRCHLQGLAKGTEVSVGLRLVHNEFFRRAKFKSLTV 1084
Qy 1052 VSTABILFNDVSFTLLPGQAFVRQOTETKVEPFPVNPPLIVGSSVGGLLLLALITAA 1111
Db 1085 VSTFELCTEESVLQLTASRWSESLLEV-VQTRPILISLWILIGSVLGGLLLLALIVFC 1143
Qy 1112 LYKLGFF-----KQYK 1123
Db 1144 LWKLGFFAHKKIPEEKREEK 1164

Search completed: November 25, 2003, 14:17:26
Job time : 14.1742 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	4358	74.1	1151	11	Q9JT30
2	3878.5	66.0	1036	11	Q8ca73 mus musculus
3	3786.5	64.4	920	6	Q28984
4	3457	58.8	1169	4	Q8iva6 homo sapien
5	3297.5	56.1	1169	11	Q9QXHA
6	3225.5	54.9	1161	11	Q9QYEV
7	1512.5	25.7	1161	11	Q9WTV4
8	1502	25.5	1160	11	Q9R200
9	1382	23.5	1196	13	Q98TF1
10	1341.5	22.8	1086	4	Q56HB1
11	1329.5	22.6	1187	13	Q98TF0
12	1264	21.5	927	6	Q8HXV0
13	1142.5	19.4	1167	11	Q88340
14	1097	18.7	1167	11	Q88341
15	1034	17.6	1171	13	Q42094
16	1020	17.3	1038	11	Q8BS01
					mus musculus
					Q9ji30 rattus norv
					Q8ca73 mus musculus
					Q28984 sus scrofa
					Q8iva6 homo sapien
					Q9qxh4 mus musculus
					Q9qyev rattus norv
					Q9wtv4 mus musculus
					Q9t200 mus musculus
					Q98tf1 cyprinus ca
					Q96hbl homo sapien
					Q98tf0 cyprinus ca
					Q8hxv0 bos taurus
					Q88340 rattus norv
					Q88341 rattus norv
					Q42094 gallus gall
					Q8bs01 mus musculus

Db 17 FNLDTENPMTFQENASFGQSVIQLGETRVVVAAPQEVKAVNQTGALYQCDYSTNRCDPI 76

```
Qy 61 RLQVPEAVNMISGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 120
Db 77 PLQVPPAVNMISGLSLAATTVPPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQKSKTILFSLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQKSKTILFSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGLRKVKVRELFINITGARKNAKILFL 240
Db 197 RTHFTFDFKRNPPKSHVRPIQLNGRTKTASGIRKVKVRELFINITGARKNAKILFL 256
Qy 241 TDGKFGDPLGYEDVPELDRGVIRVYVGLGDAFRSEKSOELNTVASKPPRPHVFOAN 300
Db 257 TDGKFGDPLGYEDVPELDRGVIRVYVGLGDAFRSEKSOELNTVASKPPRPHVFOAN 316
Qy 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 317 NFEALNTIRNLQLEKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GAFLYPSKDKASFINTRIDSDMDNDAYLGYASAVISRRNVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLIGAPHYETRGQGVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLIGAPHYETRGQGVSCPL 496
Qy 481 PRGORARWQCDAVLGYEGQOPWGRFGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDAVLGYEGQOPWGRFGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 600
Db 556 HGASVASISTSPHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 615
Qy 601 PVLKVKAMEPNPREAVNVECNDOVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
Db 616 PVLKVKAMEPNPREAVNVECNDOVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 675
Qy 661 YDLALDGRPHSRVAFNFTKNSRRTQVGLGTQVGLTQVGLTQVGLTQVGLTQVGLTQVGLT 720
Db 676 YDLALDGRPHSRVAFNFTKNSRRTQVGLGTQVGLTQVGLTQVGLTQVGLTQVGLT 735
Qy 721 SLVGTPLSAFNLPRVLAEDRQRLTALFPPEKNCNGNDNICODDLSITFSPMSDCLVYG 780
Db 736 TLVGEPLRSSRDLRVLAMEAQRIFTAMFPPEKNCNGNDNICODDLSITFSPMSDCLVYG 795
Qy 781 GPREFNVTVVRNDCGDSYRQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 796 DSRDFVSVTLRNDGDSYRQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 854
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 900
Db 855 QGVLKSTIWDINHPIFPANSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 914
Qy 901 QLELVPKAVYVWVTSYGUSTKYLNFPTASENTSRVQHQYQVNSLQKRSPLSLVFLVPV 960
Db 915 QLELVPKAVYVWVTSYGUSTKYLNFPTASENTSRVQHQYQVNSLQKRSPLSLVFLVPV 974
Qy 961 RLNQTIVNDRPOVTFSENLSSTCHTKERLPSHSDFLAEIRKAPVNCVSIACVORTQCDIP 1020
Db 975 QINKVTIWDPPQVTFSENLSSTCHTKERLPSHSDFLAEIRKAPVNCVSIACVORTQCDIP 1034
Qy 1021 PFGIQEENATLKGNSLDWYIKTSHNHLIIIVSTAEIIFNDSVFTLLPGQAGFVRSQTE 1080
Db 1035 SPNSKEIENVTLQGNLLFDWYIETSHDHLIIIVSTAEIIFNDSVFTLLPGQAGFVRSQTE 1094
Qy 1081 KVEPEVNPPLIVGSSVGGILLALITAAIYKLGFFKQYKQNMSSGGPPGAEPPQ 1137
Db 1095 KVEPEVNPPLIVGSSVGGILLALITAAIYKLGFFKQYKQNMSSGGPPGAEPPQ 1151
```

```
RESULT 2
Q8CA73 PRELIMINARY; PRT; 1036 AA.
ID Q8CA73
AC Q8CA73;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Integrin alpha M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RL EMBL; AK039444; BAC30350.1; --
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E1B8E77 CRC64;

Query Match 66.0%; Score 3878.5; DB 11; Length 1036;
Best Local Similarity 65.6%; Pred. No. 6,4e-279;
Matches 747; Conservative 126; Mismatches 146; Indels 119; Gaps 2;

Qy 1 FNLDTENAMTFOBNARFGQSVVOLQGSRRVVGAPQIIVANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFOBNARFGQSVVOLQGSRRVVGAPQIIVANORGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNMISGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 120
Db 77 PLQVPPAVNMISGLSLAATTVPPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQKSKTILFSLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQKSKTILFSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNSRLIKPIITOLLGRTHATGLRKVKVRELFINITGARKNAKILFL 240
Db 197 RIHFTFDFKRNPPKSHVRPIQLNGRTKTASGIRKVKVRELFINITGARKNAKILFL 256
Qy 241 TDGKFGDPLGYEDVPELDRGVIRVYVGLGDAFRSEKSOELNTVASKPPRPHVFOAN 300
Db 257 TDGKFGDPLGYEDVPELDRGVIRVYVGLGDAFRSEKSOELNTVASKPPRPHVFOAN 316
Qy 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 317 NFEALNTIRNLQLEKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGYASAVISRRNVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDVLIGAPHYETRGQGVSCPL 480
Db 437 ENFTGPHSTSIK----- 450
Qy 481 PRGORARWQCDAVLGYEGQOPWGRFGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 451 ----- 450
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 600
Db 451 -----SQRIIGAHFSPGLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLSQ 498
Qy 601 PVLKVKAMEPNPREAVNVECNDOVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
Db 499 PVLKVKAMEPNPREAVNVECNDOVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 558
```

```
QY 661 YDLALDSGRPHSAVNEKNTSTRROTQVULGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720
DB 559 YDLALDPGRSIRAFDETKNNTRRRRTQVFLGMLQKCEETLKLILPDCVDSVSPILRLNY 618
QY 721 SLVGTPLSAFNGLRPVLAADAQRLFTALPFPEKNCNDNICDDLSITFSFMSLDCLVVG 780
DB 619 TLVGEPLRSFGNLRPVLAADAQRFMTAMPFPEKNCNDNICDDLSITFSFMSLDCLVVG 678
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRL-ACESASSTE 839
DB 679 GPQDFNMSVTLRNDGEDSYGTQVTVYPSGLSYRKDSASQNPULKPFVFKPAESSSSSE 738
QY 840 VSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKLVNNTSENMPRINKTE 899
DB 739 GHGALKSTWNINHPIFPANSEVTNITFDVDSHASFGNKLKLVNNTSENMPRINKTE 798
QY 900 FOLELPVKYAVYVVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVP 959
DB 799 FOLELPVKYAVYVVTSDRESSIRYLNFTASEMTSKVIHQYQPNLQGRSLPVSUVFWIP 958
QY 960 VRLNQTIVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCRIQCDI 1019
DB 859 VQINNVTVDHPQVIFESQNLSSACHTEQKSPHSNFRDQLERTPVLNCSVAVCKRIQCDL 918
QY 1020 PFGIOQEFNATLKGNSLFDWYIKTSHNHLIIIVSTABIIFNDSVFTLLPQCGAFVRSQTE 1079
DB 919 PSPTQEIENVTILKGNLSDFWYIKTSHGHLIIIVSTABIIFNDSVFTLLPQCGAFVRSQTE 978
QY 1080 TKVEPEVNPPLIIVGSSVGGILLALITALYALYKLGPKRQYKQNMWSSGGPGASPO 1137
DB 979 TKVEPEVNPPLIIVGSSVGGILLALITALYALYKLGPKRQYKQNMWSSGGPGASPO 1036

RESULT 3
Q28984
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CD11b (Fragment).
GN CD11b.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWF_A; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

```
Query Match 64.4%; Score 3786.5; DB 6; Length 920;
Best Local Similarity 78.4%; Pred. No. 3.6e-272;
Matches 722; Conservative 85; Mismatches 113; Indels 1; Gaps 1;

QY 118 POKFPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMQYS 177
DB 1 POKFPEALRGCPQEDSDIAFLIDSGSIINRLDFQRMKEFVSTMVGMQFQKSKTLFALMQYS 60
```

```
QY 178 EBFRIHTFKERQNNPNRSLIKPITOLLGRTHRTATGLRKVYRELPNITNGARKQNAFKIL 237
DB 61 EDFYHTFTFNDPKRNPSPKLLVRPILQLGRTHRTATGIRKVVYRELPHSKSGARENAFKIL 120
QY 238 FLLTGEKFGDPLGVEDVITPELDRGVIRYVILGFGDAFPESEKSRQELNNTVASKPRPDHVF 297
DB 121 VVITGEKFGDPLGVEDVITPEADRGVIRYVILGFGDAFNSKWSRELNNTIASPCGDHVF 180
QY 298 QANNEALKTQVQNLREKIFAIEGTQTSSTSSSEHEMSQEGFSAAITSNGLLSTVGSVD 357
DB 181 QVNFNEAVKTIQNLQEKTFATIEGTQTSSTSSSEHEMSQEGFSAAITSNGLLSTVGSVD 240
QY 358 WAGGVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVLCAPRYQHIGLVA 417
DB 241 WAGGAPLHPKPDVFIINTTRVDSMDNDAYLGYAVEVILNRQAQSLVLCAPRYQHTGLVY 300
QY 418 MFRQNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYVEOTRGQOVSV 477
DB 301 MFKQNSGAWEKWADIKGSGIYFGASLCSVDVDRGSDVLVLIGAPHYVEOTRGQOVSV 360
QY 478 CPLPRGORARMQCDVAVLYGEGQOPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEENRGAV 537
DB 361 CPLPQG-RAKWCQCVILCGEQGHPSRFGAALTALGDVNGDKLTDVAIGAPGEEENRGAV 419
QY 538 YLPHGTSGSGISPSHSQRIAGSKLSPRIQYFCQSLSGGQDLTMDGLVDLTVGAQGHVLL 597
DB 420 YLPHGTSELGISPSHSQRIAGSKLSPRIQYFCQSLSGGQDLTMDGLVDLTVGAQGHVLL 479
QY 598 RSQPLRVKAIEMEPREVARNVFECNDQVVKGEAGEVRLVHVKSTDRDLREGQIOS 657
DB 480 RSQPLRVKAIEMEPREVARNVFECNDQVVKGEAGEVRLVHVKSTDRDLREGQIOS 539
QY 658 VVYDIALDSGRPHSAVNEKNTSTRROTQVULGLTQTCETLKLQLPNCIEDPVSPIVLRL 717
DB 540 IITYDALDPGRPHPRVAFVEETKNNTRRQTLGLSRKCEHLALWLPDCVEDSVTPVLR 599
QY 718 LNPSLVGTPLSAFGNLRPVLAEDAQRLFTALPFPEKNCNDNICDDLSITFSFMSLDC 777
DB 600 LNPSLVGTPASSFGNLRPVLAEDAQRLFTALPFPEKNCNDNICDDLSITFSFMSLDC 659
QY 778 VVGGRPEFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESAS 837
DB 660 VVGGRPEFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESAS 719
QY 838 TEVSALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKLVNNTSENMPRINK 897
DB 720 TEVSALKSTSCSINHPIFPENSEVTNITFDVDPDAFLGYKLLKLVNNTSENMPSSNK 779
QY 898 TEFQLELPVKYAVYVVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLPISLVFL 957
DB 780 TEFQLELPVKYAVYVVTSLVSTKYFNFTASEKTRHVIHQYQFNGLQGRKLPISVWF 839
QY 958 VVRLNQTIVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCRI 1017
DB 840 VVRLNQTIVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCRI 899
QY 1018 DIPFGIOQEFNATLKGNSL 1038
DB 900 DIPFGIOQEFNATLKGNSL 920

RESULT 4
Q81VA6
ID Q81VA6 PRELIMINARY; PRT; 1169 AA.
AC Q81VA6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

RN PP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038237; AAH38237.1;
SQ SEQUENCE 1169 AA; 128521 MW; A17B484FEFC79EB6 CRC64;

Query Match 58.8%; Score 3457; DB 4; Length 1169;
Best Local Similarity 60.7%; Pred. No. 1.7e-247;
Matches 685; Conservative 143; Mismatches 295; Indels 6; Gaps 4;

Qy 1 FNLDTENAMTFOENARFGQSVVQVANSWVVGAPQKITAANQTGGIYQCYSTGACEPI 60
Db 20 FNLDTTELTAFRVDAGFGDSVVQVANSWVVGAPQKITAANQTGGIYQCYSTGACEPI 79

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 120
Db 80 GLQVPPPEAVNMSLGLSLAATTSPPQLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 137

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRPMKELVSTIMEQKKSTLPSLMQYSEEP 180
Db 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFMNFVRAVISQFQPSQFSLMQFSNKF 197

Qy 181 RIHFTFEFQNNPNSRLIKPIITQLGRTHATGLRKVVRVRELNITNGARKNAFKILFL 240
Db 198 QTHFTFEFRSSNPLSLASVHQLOQFTYATATQNVVHRLFHASGARRDAAKILIVI 257

Qy 241 TDGKFGDPLGVEDYIPELDREGVIRYVVGFGDAFRSEKSRQELNTVASKPRDRHVFQAN 300
Db 258 TDGKFGDPLGVEDYIPELDREGVIRYVVGFGDAFRSEKSRQELNTVASKPRDRHVFQAN 317

Qy 301 NPEALKTVQNLREKIFAIEGTQGTSSSPHEMSQEGSAITNSGPLLSTVGSYDWAG 360
Db 318 DFDALKIQONLQKREKIFAIEGTQGTSSSPHEMSQEGSAITNSGPLLSTVGSYDWAG 377

Qy 361 GVFLTSKEKSTFTINMTVDSDMDAYLGVAIAIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFTINMSQVMDRSLYCYSTELALWKGVOSLVLGAPRYOHTGKAVIFT 437

Qy 421 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTRGGVSCPL 480
Db 438 QVSRQWRMKAETGTQIGSYFGASLCSVDVDSNGSTDVLVLGAPHYEQTRGGVSCPL 497

Qy 481 PRGARQWOCDAVLYGEGQWPGFRGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVLYF 540
Db 498 PRGWR-RWMCDAVLYGEGQWPGFRGAALTVLGDVNGDKLTDVWIGAPGEKENGAVLYF 556

Qy 541 HCTSGSGISPSHSORIASKLSPLQYFGOSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGVLGPSISPSHSORIASKLSPLQYFGOSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

Qy 601 PVLRYKAIMFNPVEAVNFCNDQVVKGEAGEVRYVCLHVOKSTRDRLRREGQTQSVYT 660
Db 617 PVLWVGVSQMOTFPAEIPRSAFECEQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

Qy 661 YDLALDSRPHSRVFNETHQSTRQTVGLGTQTCETLKLQPLNCIEDPSPVLRNLP 720
Db 677 LDALDPGRLSPRATFQETKNSLRVRLGLKAHCENFNLLPSCVEDSVTPITRLNF 736

Qy 721 SLVGTPLSAGNLRVLAEDAORLTALFPPEKNCNDNI CODDLSTIFSMSLDCLVVG 780
Db 737 TLVGLPLAFNRLRPLAADAQRYFTASLPPEKNCNDNI CODDLSTIFSMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDEGDSYRTQVTEFFPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 SNLELNAEVMWVWNGEDSYGTTVTFSHPAGLSYRYVAEGQKQGLRSLHLATCDSPVGV- 854

Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db 855 SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVLGDRLILLTANVSENNTPTRTSKTF 914

Qy 901 QLELPVKYAVYTVVSSHEQFTKYNLNFSESEKESHVAMHYQVNNLQORDLPVSNFWMP 974
Db 915 QLELPVKYAVYTVVSSHEQFTKYNLNFSESEKESHVAMHYQVNNLQORDLPVSNFWMP 974

Qy 960 VRLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAQPVNVCISIAVCORIQCDI 1019
Db 975 VELNQEAVMDVEVSHPPQNSLRCSSEKIAFPASDFLAHIQKNPVLDCSIAGCLRFPCDV 1034

Qy 1020 PFGIOEEFNATLGNLSPDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQOGAFVRSQTE 1079
Db 1035 PFSVQOEELDTLGNLSPDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQOGAFVRSQTE 1094

Qy 1080 TKVEPFEVNPPLIIVGSSVGGLLLALITALYALYKLGFFKQYKDMWSE 1128
Db 1095 TVLEKYKHNPTELIIVGSSVGGLLLALITALYALYKLGFFKQYKDMWSE 1143

RESULT 5
Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
AT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DR 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RA Huang X., Goreki K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA Tsuchiya H.
RA "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF211864; AAP23492.1;
DR HSP; P11215; IBHQ.
DR MGD; MGI:96609; ItgaX.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWA; 1.
KW Integrin.
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match 56.1%; Score 3297.5; DB 11; Length 1169;
Best Local Similarity 56.2%; Pred. No. 1.2e-235;
Matches 641; Conservative 174; Mismatches 306; Indels 19; Gaps 7;

Qy 1 FNLDTENAMTFOENARFGQSVVQVANSWVVGAPQKITAANQTGGIYQCYSTGACEPI 60
Db 20 FNLDAEKLTHFMDGAEFGHSGVLYQDSSVWVVGAPKEIKATNQIGGLYKCGYHTGACEPI 79

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 120
Db 80 SLQVPPPEAVNMSLGLSLAATTSPPQLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 138

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRPMKELVSTIMEQKKSTLPSLMQYSEEP 180
Db 139 FPTAQECPRQEQDIVFLIDGSGSISSRNFMNFVRAVISQFQPSQFSLMQFSNKF 198

Qy 181 RIHFTFEFQNNPNSRLIKPIITQLGRTHATGLRKVVRVRELNITNGARKNAFKILFL 240
Db 199 RVHFTFNFNISTSPSLSGVRLQRYGTYTASAIKHVITELFTTQSGARQDQATKVLIVI 258

Qy 241 TDGKFGDPLGVEDYIPELDREGVIRYVVGFGDAFRSEKSRQELNTVASKPRDRHVFQAN 300
```

```
Db 259 TDGRKQDNLSYSDVIMAEAAIIIRYAIAGVGFYNEHSGKELKAIASMPSEHYVFSVE 318
Qy 301 NFALKTVQNLREKIFAIEGTOTIGSSSSFEHMSQEGFSAATISNGPLLSVTGSDWAG 360
Db 319 NFALKDIEQNLREKIFAIEGTETPSSSTFEHMSQEGFSAVTPDGPVLGAVGFSWSG 378
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNVRQSLVLGAPRYQHTGLVAMFR 420
Db 379 GAFLYPSNMRPTFINNSOENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIFT 438
Qy 421 QNTGMWESNANVKTGTOIGYFAGSLCSVDVDSNGSDTLVLIGAPHYYEQTGRGQVSVCLP 480
Db 439 QESRHRPKSEVRGTIGSYFAGSLCSVDMDRGGSTDLVLIGVPHYYEHTRGQVSVCPM 498
Qy 481 PRGORARWQCDVLYGQOGPWGRFGAALTVLGDNVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 499 P-GVGRWHCGTTLHGEGHPWRFGAALTVLGDNVNGDSLADVAIGAPGEENRGAVYIF 557
Qy 541 HGTSGGISPSHSQRIAGSKSLSPRLQYFGOSLGGQDLTMDGLVDLTGAGQHVLRLRSQ 600
Db 558 HGASRQDIAPSPORISASQIPSRIOYFGOSLGGQDLTRDGLVDLAVGSKGRVLLRTR 617
Qy 601 PVLVRKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDRLRREGIOISVVT 660
Db 618 PILRVSPVTHFTPAEISRSYFECQEVAPRQTLSDATVCLIHIESPKTQL--GDLRSTVT 675
Qy 661 YDLALDSGRPHSAVNETKNSRRTQVGLTGTCTETLKLQLPNCIEDPVSIVLRLNF 720
Db 676 FDALDHGRUSTRRAIFKETKTRALTRVKTGLNKHCSVKLLLPACVEDSVTPITLRLNF 735
Qy 721 SLVGTPLSARGNRPVLAEDAQRALTALPFFKXNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 SLVGVPISSQLNQLPMLAVDDQYFTASLPFFKXNCGADHICQDDLSVFGFPDLKTLVG 795
Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDSVRKYSTLQ-----NORSQSWR 829
Db 796 SDLELVNVTVNSDGEDSYGTTVTLFYPVGLSPRRVAEQVFLRKXEDQOMQRGQSHLH 855
Qy 830 LACESASSTEVGALKSTSCSINHPIPPENSEVTENITFDVDSKASGNKLLKANVTSE 889
Db 856 LMCDD--STPDRSGLSTSCSRHVIFRGSGQMTFLTFDVSFPAELGRLRLRARVGE 913
Qy 890 NMPRTNKTFFQLELPKVAVMVMTSHGVSTKYLNFPTASE--NTSRVMOHOYVSNLGR 948
Db 914 NNVPKPTTFFQLELPKVAVMYTWISSHDQFTKYNLFSSEKTSVVEHFRFVNILGR 973
Qy 949 SLPLSVLVPVRLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCS 1008
Db 974 DVPVSINFVPIELKGEAVM-TVMVSHQPPLTCQYRNRLKPTQFDLLTHMQKSPVLDCS 1032
Qy 1009 IAVCQRIQCDIPFGIQEENATLKGNSLFDWIKTSHNHLILVSTAELFNDVSFTLLP 1068
Db 1033 IADCLHLRCDIPSLGIDELFYFLKGNLSFGMTSQTLOKVKLLLSAEITFNFSVSQLP 1092
Qy 1069 GQAFVRSOTETKVEPPEVNPPLVLIVGSSVGLLLALITALYKLGFFKROYKDMSE 1128
Db 1093 GQEAFLRAQTKTVLEMYKUNPVLIVGSSVGLLLALITALYKAGFFKROYKEMLEE 1152

RESULT 6
Q9QVE7
ID Q9QVE7 PRELIMINARY; PRT; 1161 AA.
AC Q9QVE7;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietech G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin.
SQ
SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 54.9%; Score 3225.5; DB 11; Length 1161;
Best Local Similarity 57.2%; Pred. No. 2.6e-230;
Matches 645; Conservative 163; Mismatches 306; Indels 13; Gaps 9;

Qy 2 NLDTNAMTFQENARGFGOSVVQLOGSRVVVCGAPQEIIVAANQSGSLYQCDYSTGSCPTIR 61
Db 21 NLDVEEPIVFRDEDAASFGQTVVQFGSRLVVGAPLEAVAVNTQTRLYDCAPATGMCOPIV 80
Qy 62 LQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSENTYVVKGLCFLFGSNLRQOPQKF 121
Db 81 LASPLEAVNMSLGLSLVTATNNAQLLACGPTAQRACVKNMYAKGSCLLLGSSL-QPIQAV 139
Qy 122 PVALRGCPQEDSDIAFLIDGSGSIIPHDPFRKVELVSTIMEQKSKTLFSLMQSEBEPR 181
Db 140 PASMEPCERQEMDIAFLIDGSGSINQRDPAQMKDFVKALMGFASTSTLFLSMQYSNLIK 199
Qy 182 IHFTKEFONNPNPSLIKPIQLGRTHATGLKRVVRELFNITNGARKNAFKILFLIT 241
Db 200 THFTTEFNILDDPSLDVPIVQLQGLTYTATGIRTVMEELPHSKNGSKRSKAKILLVIT 259
Qy 242 DGEKFGDPLGVEDVPELDEGVIRYVLFQGFDAFRSEKSRQELNTVASKPPRDHVFQANN 301
Db 260 DQKYRDPLEYSDVIPAADKAGIIRYAGVGDAFOEPTALKELNTIGSAPPQDHVKVGN 319
Qy 302 FEAATVQNQLREKIFAIEGTQTSSTSSFEHMSQEGFSAATISNGPLLSVTGSDWAG 361
Db 320 FAALRSIORQLEKIFAIEGTQTSSTSSFEHMSQEGFSAATISNGPLLSVTGSDWAG 379
Qy 362 VELYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNVRQSLVLGAPRYQHTGLVAMFR 421
Db 380 AFLYPNTRPTFINNSOENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIFT 439
Qy 422 NTGMWESNANVKTGTOIGYFAGSLCSVDVDSNGSDTLVLIGAPHYYEQTGRGQVSVCLP 481
Db 440 EAHHRPKSEVRGTIGSYFAGSLCSVDVDRDGSSTDLVLIGAPHYYEQTGRGQVSVFPV 499
Qy 482 RQORARWQCDVLYGQOGPWGRFGAALTVLGDNVNGDKLTDVAIGAPGEDNRGAVYLF 541
Db 500 -GVGRWQCEATLHGEGHPWRFGAALTVLGDNVNGDNLADVAIGAPGEESRGAVYIF 558
Qy 542 GTSGSGISPSHSQRIAGSKSLSPRLQYFGOSLGGQDLTMDGLVDLTGAGQHVLRLRSQ 601
Db 559 GASRLTEIMPSPQRVTGSQLSLRLQYFGOSLGGQDLTQDGLVDLAVGAGQHVLRLSLP 618
Qy 602 VLVRKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDRLRREGIOISVVT 661
Db 619 LLKVELSTRFAPMEVAKAVYQCEWETPTVLEAGEATVCLTVHKGSPDLL--GNVQSVRY 676
Qy 662 DIALDSGRPHSAVNETKNSRRTQVGLTGTCTETLKLQLPNCIEDPVSIVLRLNFS 721
Db 677 DIALDPGRILSRAIFDETKNCTLTGRKTGLGDHCETVKLLPDCVEDAVSPILRLNFS 736
```


Db 118 SYKTEFDSDYVVRKDDPALLKHVXKMLLLTNTFGAINVATEVFREEGARPDPATKVL 177
Qy 239 LITDGEKFGDPLGYEDVIFELDREGVIRYVLFGDAPRSEKSRQELNTVASKPPRDHVFQ 298
Db 178 IITDGE--ATDSNIDAAND---IIRVIIGIKHFQTKESQETLHKPASKPASEFVKI 230
Qy 299 ANNFEALKTVONOLREKIFAIEGTQTGSSSPHEMSQEGFSAITNSGPLLSTVGSYDW 358
Db 231 LDTFEKLKOLFTLOKIVIEGTQKDLTTFNMELSSGISADLSRGHVVAVGAKDW 290
Qy 359 AGGVF-LYTSKESKSTFINTRVDSMDNDAYLGA---IILNRVQSLVLGAPRYOHIGLV 416
Db 291 AGFGLDKADLQDDTFIGHEPLTPEVRAGLVGTVTWLPSPRQKTSLLASGAPRYQMGV 350
Qy 417 AMFR--QNTGMWESNANVKTGTFAGFASLCSVDVDSNGSTDVLIGAPHYEQTRGGQ 474
Db 351 LLFQEPQGGHWSQVTHGTQIGSYFGELCGVDVDQGETELLIGAPLFYGEQGR 410
Qy 475 VSCPLPRGRARQCDV--LYGEOGPWGRGAAITVLGVNGDKLTDVAIGAPGED 532
Db 411 VFYI----QRQLGFEEYSELQDQPGYPLGRFGEAITALTDINGDLVDVAVGAPLEE- 464
Qy 533 NRCAYLHGTSGSGISPSHSQRIAGSKLSPRQYFGOSLSSGQDGLTMDGLVLTVAQS 592
Db 465 -QNAVYFNGRHG-GUSPQSQRIEQTQVLSGIQWFGRSIHGVDKLEGDLADVAGAS 522
Qy 593 HVLLRSQPLVRKALMEFNPREVARNFECNDQV--KGKEAGEVRVCLHVKQSTRDLR 651
Db 523 QMIVLSRPPVDVMTLMSFPAIPVHEVECSYSTSKMKGVNITICQI-KSLIPQF- 580
Qy 652 EGOIQSVVYDIALDGRPHSRVAFNETKNSTRQTOVLGLTQCTETLKLQIPNCIEDPV 711
Db 581 QGRVANLTYTLQDGHRTRRRGLFGRHRLRNIAVT-TSMSCDTSFSPHPCVQDLI 639
Qy 712 SPVLVLSNLSL---VCTPLSAPGN-----LRPLAEDAQRPLTALFPREKACGNDNIQ 762
Db 640 SPINVSLSNLSLWEEBETPDQRAGKDIPIPLRSLSHSETWEI-----PFEKNGEDKKCE 694
Qy 763 DDLISITFSMSLDCVLVGGPREFNVTVRNDCGDSYRTQVTFPFLDLISYRKVSTLQ 822
Db 695 ANLRVSFSPARSALALTAFASLSVELSLNLEEDAYWQLDHFPPGLSFRKVELM--- 751
Qy 823 RSORSRLACES--ASSTVSGALKSTSCINHIPIIPENSEVTNITFDVDSKASLGNKL 880
Db 752 KPHSQIPVSCIELPEESRLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNSSWGDV 808
Qy 881 LKANYTSEN---NPRINKTEFQLELPVKYAVMVVTSYSHGVSTKYLNFTASENTRVM 936
Db 809 ELHANYTCNNEDSDLLSDNSATTI---IPILYPINILIQDQEDSTLYVSFTPKGPKHQV 865
Qy 937 QHOYQV---SNLQORSIP-LSLVLVPLVRLNQTVWDRPOVTFSENLSSTCHTK--ERLP 990
Db 866 KHWYQVRIQPSIHDHNIPTLEAVVGVPOPPSEGPITHQWSVQMEPPV--PCHYEDLERLP 923
Qy 991 SHSD--FLAELRKAPVNVCSIAVCORIQDIPFFGIQEEFNATLKNLSFDWYIKTSHNH 1048
Db 924 DAAEPCPLGALPCPVV-----FRQELVQVIGTLELVGIEAS-SM 964
Qy 1049 LLIVSTABILFNDVSTLLPQCAFVRSQTETKVEFEFVNPPLPLVGVSGVGLLLALI 1108
Db 965 FSLCSLSISFNSKHFHLYGSNASL-AQVVMKVVDVYVYKQMLYLVLSGIGGLLLLI 1023
Qy 1109 TAALYKLGFPFKROYKDMWSEG-GPPGAEP 1136
Db 1024 FIVLYKVGFPKRLKEMKEMAGRVNPIP 1052

RESULT 11

Q98TF0

ID Q98TF0 PRELIMINARY; PRT; 1187 AA.

AC Q98TF0;

DT 01-JUN-2001 (T-EMBLrel. 17, Created)

DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE CD11-2.
GN CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048537; BAB39135.1; -;
DR HSSP; P20701; 1LFA.
DR InterPro; IPR001969; Asprotease site.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;

Query Match 22.6%; Score 1329.5; DB 13; Length 1187;
Best Local Similarity 31.0%; Pred. No. 2.2e-89;
Matches 369; Conservative 210; Mismatches 456; Indels 157; Gaps 43;

Qy 1 FNLDTENAMTFOENARG-FGOSVQLO-GSR--VVVGAPOEIVAAQORSLYQCDYSTGS 56
Db 32 FNIDTEHPURNGTDEDFGYSVYQTEFGNRKQIIVGAPLE---GNSAGEMYSTADLOS 88
Qy 57 CEPRIQLQVP---VEAVNMSLGLSLAATTPSPQLLACGPTVHTQTSNTYVVKGLCFPLGS 112
Db 89 CK--RLRPGSSRVFFGMSAAVSSAULTS-----CSPYFAHECDGNSYLVNGVCYQFNS 140
Qy 113 NLRQOPQKPEALRQCPQEDSDIAFLIDGSGSIIIPDPRMKELVSTIMEOLKSKTLPFS 172
Db 141 SL-QAVSNFTAAAYQCSKREVNLFVLFDFGSSSMKAVEFDMKNFIDKVMKLSNSSIKFA 199
Qy 173 LMOYSEERIHTFEKFNQNPRLI KPIITOLLGRTHTATGLRKVVRLFN-ITNGARK 231
Db 200 AVQFSTEIRTVDFDNDYQNGSAEELMKE-RHMKSLTNTYKAINYYLVKNVLSVSGADP 258
Qy 232 NAFKILFLITDCEKFGDPLGYED--VIPELDREGVIRYVLFGDAPRSEKSRQELNTVAS 289
Db 259 NAKALVIITD---GDSDDNDYLNLCDEQNILRYIIVG-----KVDLTTLTQLAA 309
Qy 290 KPPRDHVFOANFEALKTVQNOLREKIFAIEGTQTGSSSPHEMSQEGFSAITNSGPL 349
Db 310 EPKLANTFYIQSYNGLKGLLDNLQKIVNIEGSKAHRDQKELSSQSGFVVYQESVI 369
Qy 350 LSTVGSYDWAGVFLYT---SKEKSTFFINMTVDSMDNDAYLGYAAAIIILNRVQSLVLG 406
Db 370 VGSVGSNDWRGALYEVMSGSKFRQTEITDPVAVN---KDSYMGYSTVLGMRHGVSLLFG 426
Qy 407 APRYOHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHY 466
Db 427 APRAEHTGLVLTFTKNESTWTVMRNINGEIQISYFGASLSDVDSDGSDFLVLVAPLP 486
Qy 467 YE-QTRG-QQSVCPPLRGQRARWQCDVLYGEQ--GQPMGRFGAALTVLGVNGDKLTD 522
Db 487 YSQPRAEGRLYVYTL-----SEQYSQKTLQSTTGRFATSLASLKDINGDGLSD 535
Qy 523 VAIGAPGEDNRGAVLPHGTSGSGISPSHS--ORTAGSKLSPLRQYFGOSLQSGQDLTWD 581
Db 536 VAVGAPLE--NEGVMVYILGDGTHGINPEHAPQIPARSVLFGLOQFGVSLSGQMDMND 593

```
Qy 582 GLVLTGVAQGHVLLRLSQPVLRVKAIMEFNPREVARNVFECNDQVVGKEAGVRVCLH 641
Db 594 NLPDIVIGTGGVILLNARPVMSQAQLSPNPMISLNYFPCPGS--NAFNAPNLTSCT 651
Qy 642 VOKSTRDLRGQIQSV--VTYDIALDSGRPHSRVAFNETKXSTR--RQTVGLGTQTC 697
Db 652 VTERSS--TGSLEKUNVSLNMLDVRGMSRGFFDPMDSSTRTQQSVLLDSGSCS 708
Qy 698 TLKLOLPNCIEDPVSPVILRLNES----LVCTPLSFAFGLNLPVLAEDAORLFTALFPPEK 753
Db 709 NFSFMLRCVADTVSPKIRNFMSOTOMLSGNSLAUL--DIQSRTEEYEVUL-----FOR 761
Qy 754 NCGNDNICQDDLSITFSFMSLCLVVGGRPREFNVTVVRNDGEDSYRTQVTFPPDLDSY 813
Db 762 NC-NSNSCVADLKLNFSTN-DTLVVENQAHTVLVSLANFGDDSYNTSIVLHYPEGLSL 819
Qy 814 RKVSTLQNRORSWRLACESASSTEVSGALKSTSCSINHIFPENSEVETNITFDV--- 870
Db 820 SKFDAIKPSRTR-----SSCGDRDSGATNRITCSINLPVYRSGITTTQFLGFRVTKW 871
Qy 871 DSKASISGNKLLKANVTSENNMPTNKTEFOLELPVKYAVVMVYTHGV-STKYLNFAS 929
Db 872 DYVSDRMEMITANSNDNGNN---SDMSVRSRIPVQFAVELAISLVAEDSVTYLNFSL 928
Qy 930 ENTSRVMQHQVSNLQSRSLPISLVPLVRLNQTIVWDRPQVTFSENLSSTCHTKERL 989
Db 929 DRGPKPLNIYKVNVLKGLPVSVTLSLPCQ-----TTHVLTTHNFMSQ 974
Qy 990 PSHSDFLAELRKAPVVC-----SIAVCQIQCDIPFGIOEE 1027
Db 975 EVHVSFISSYHQ--IIMCLLNKHLFPSPELSAVQTRTTGRSLWC-----VSSITSGEI 1026
Qy 1028 FNATLKGNL-----SFDWIKTSHNLLIVSTAEILFNDSVFTLLPG-QGA 1072
Db 1027 FRSSV--NLMAEAVLNQVKEYESKYSFY-EFRDHFVNIS-AELNFTSRNQSTGLKYN 1082
Qy 1073 FVRSOTETKVEPFEPNPLPLVSGVGLLALITALYALYKLGFFKQYKD 1124
Db 1083 PHRSQTEKVEFVPPSRMLIVCTGAVGGFFLLIILULLKCGFFKRNRPD 1134

RESULT 12
ID Q8HZV0 PRELIMINARY; PRT; 927 AA.
AC Q8HZV0;
DT 01-MAR-2003 (TrEMBLrel. 23; Created)
DT 01-MAR-2003 (TrEMBLrel. 23; Last sequence update)
DE Lymphocyte function-associated antigen 1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1];
RP SEQUENCE FROM N.A.
RA Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Sequence of the alpha subunit of bovine lymphocyte function-
associated antigen 1."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF40778; AAN63636.1; --
FT NON_TER 1 1
FT NON_TER 927 927
SQ SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;
```

```
Query Match 21.5%; Score 1264; DB 6; Length 927;
Best Local Similarity 34.7%; Pred. No. 1.1e-94;
Matches 339; Conservative 165; Mismatches 388; Indels 86; Gaps 28;

Qy 174 MOYSEFRHTFTKFEQNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNA 233
Db 4 VQSTYFRTFTPLDIQKDPDALLAGVKHMLLNTTFGAINVYAKEVFRPDLGAPDA 63
```

```
Qy 234 FKILFLLTQGEKFGDPLGYEDVIPELDRQGVIRYVLGFGDAPRSEKSRQELNTVASKPPR 293
Db 64 TKVLIITITDGK---PPTNTTLMRPKTS-----RSLIGKGNFKTKESQEAHQFASKPVE 116
Qy 294 DHYFOANNFEALQVQNOLREKIFAJEGTQTGSSSSFEHEMSQEGFSAAITSGNPLSTV 353
Db 117 EFVKILDTFELKLDLTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSEHGUVGAV 176
Qy 354 GSYDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGYAAA-IILNRNVQSLVLGAPRYQ 411
Db 177 GAKDWAGGFLLKADLKSSITFVGNELPITVESRAGLYGTVTRLPSSRGTMSELLATGAPKYQ 236
Qy 412 HIGLVAMFRO--NTGMWESNANVKGTQIGAYFAGSILCSVDVDSNGSTDLVLICAPHYEQ 469
Db 237 HVGRVLLFQPKRGPGMSQIQEIDGIGSYFGELCGVDVDRDGETELLIAAPLYGE 296
Qy 470 TRGOVSVCPLPGRQARMOCDAVLGEOQOPGRFGAALTVLGDVNGDKLTDVAICAPG 529
Db 297 ORGRGVFIY---QKIQLFQWSELQGETGYPLGRFGAAIAALTIDNGDELTDAVAGPL 353
Qy 530 EEDNRGAVYLFHGTSGSGISPSHRSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLDVTVG 589
Db 354 EE--QGAVYIFNGQOG-GLSPRPSQRIEGTOMFSGIQWPGRSIHGVKDLGGDGLDAVAG 410
Qy 590 AQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVGKEAG-EVRVCLHVQK--ST 646
Db 411 AEGQIVLSSRPVVDIIITSVSFPAEIPVHVEVECSYSTSNQKKEGVNLTVCQVKSLIST 470
Qy 647 RDLRREGIOISVVTYDIALDSGRPHSRVAFNETKSNRRTOTVLGLTQTCETLKLQPNC 706
Db 471 ----FOGHLVANLTYTTLQDGHRTSRGLPFGGKHKLIGNTAVTPV-KSCFVFWFHPIC 525
Qy 707 IEDPVSPIVLRNLSL---VGTPLS--AFGNLRPLVAEDAORLFTALFPPEKNCNDNIC 761
Db 526 IQDLISPINVSLSYSLWEESTPRDPRALDRDIPILKPSHLETKEIPPEKNGCEDKNC 585
Qy 762 QDDLSITFSFMSLCLVVGGRPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQW 821
Db 586 EADLKAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSPFQGLSFRKVEIL-- 643
Qy 822 QRSQSWRLACESASSTEVSGALKSTSCSINHIFPENSEVETFNITFDVDSKASLGNKLL 881
Db 644 -KPHSHVPVGCCELPEEAVHVS-RALSCNVSSPIFGEDSWDIOVMNTLQKSGWGFIE 701
Qy 882 LKANVTS-----ENNMPRTNKTEFQLELPVKYAVVMVYTHGVSTKYLNFTASENTS 933
Db 702 LQANVSCNNESSLLEDSATTS-----IPVMYPINVLTKDQENSTLYISFTPKSPRI 754
Qy 934 RVMQHQYQV-----SNLQSRSLPISLVPLVRLNQTVI---WD---RQVTFSS-ENLSST 982
Db 755 HHVKHIYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESP 813
Qy 983 CHTKERLPSHSDFLAELRKAPVVCSTIACVQICODIPFGIOEFNATLKNLSFDWVI 1042
Db 814 SDEAE-----SCSFGT--EFRCPIDF---RQILVQVNGWELRGTI 850
Qy 1043 KTSNHLILIVSTAEILFNDSVFTLLPGQAFVRSQETETKVEPFEPNPLPLIVSGSVGL 1102
Db 851 KAS-SMLSLCSLSLAISFNSSKHFLHGRNASM-AQVVMKVLDLYVEKEMLYLVLSGIGGL 908
Qy 1103 LLLALITAAALYKLGFFKR 1120
Db 909 LLLFLIFTALYKVGFFKR 926

RESULT 13
ID O88340 PRELIMINARY; PRT; 1167 AA.
AC O88340;
DT 01-NOV-1998 (TrEMBLrel. 08; Created)
DT 01-NOV-1998 (TrEMBLrel. 08; Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23; Last annotation update)
```

```

DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020045; AAC23662.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWF A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR SS SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match 19.4%; Score 1142.5; DB 11; Length 1167;
Best Local Similarity 28.7%; Pred. No. 1.7e-75;
Matches 354; Conservative 213; Mismatches 461; Indels 205; Gaps 43;

QY 1 FNLDTEA--MTQENARGFGSVQVQLGSRVVVGAPQEIIVANRGSS-----LYQCDY 52
DB 20 FNVVDVWAVTALQCPAPVLSLLHLDPSN-----NQICLLVARESSNRNTAALYRCAL 74
QY 53 STGSCPIRLQVPAEAVNNSLGLSLAATT--SPQLLAC-GPTVHGTCSNTVYKGLCLFL 109
DB 75 SI-SPDEIACQ-PVEHICNPKRGYQGVTLVGNHNGVLVCIVQVQARFRSLNSELTCACSL 132
QY 110 FGSNLRAQPKPFPEARLG-----C-----PQE 131
DB 133 LTPNLDLQAQYFSDLEGFLDPGAHVDSGDYCRSKGSGTEEEKSARRRTVEEEDDED 192
QY 132 DSDIAFLIDGSGSIIPHDFRMKELVSTIMEQL--KSKTLFSLMOYSEEFRIHFTKEFP 189
DB 193 GTEIAVLDSGSGIERSDQAKNFISTWMNRYEKCFCNCFALVQYGAIVQTEFDLQES 252
QY 190 QNNPNRSLIKPTQLLGRHTHTATGLRKVVRELFNITNGARKNAFKILFLLTDGEKFGDP 249
DB 253 RDINASLAKVQSVQVKEVTKTASAMQHVLDNIFIPSRGSRKALKVMVVLTDGDI FGDP 312
QY 250 LGVEDVIPELDRGVIRYVLGFCDAERSEKSRQELNTVASKPPRDHVFOANNPEALKTVQ 309
DB 313 LNLTTVINSFKMVGVRFAIGVDGAFKNNNTVRELKLIASDPKEAHTFKVTNYVSALDGLL 372
QY 310 NOLREXIFAIEGTQGTSSSFHEMSQEGFSAAITSNGP--LLSTVGSYDWAGGVFLY-TS 367
DB 373 SKLQQRIVHMEGT---VGDALQYLAQTGFSQAQLDKGQVLLGTGAFNWSGGALLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DMNDAYLGYAAAILNRVQSLVGLGAPRYQHIGLVAMFRQNTGM 425
DB 430 NGRGCFLNQTKEDSRSTVQSYLGYSLAVLHKAHGISYVAGAPRHKLGAVPFLRKEDE- 488
QY 426 WESNA---NVKGTQICAYGACSLCSVDVDSNGSTDVLVLGAPHYEQTGGQVSVCLPR 482
DB 489 -EEDAFVRRIEGQMGSYFGSVLCPVDIDMDGTFDGLLAAPFYHIRGEEGRVYVQVPE 547
QY 483 GQARWQCDVAVLGEQGPWGRGAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSCHPGLTNSRCGFMAAVGDIINQDKFTDVAIGAPLEGFAGDGAASYGS 606
QY 537 VYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQGSLSGGQDLTMDGLVDLTGVAQGHVLL 596

```

```

DB 607 VVIYNGHSG-GLYDPSQOIRASSVAGSLHYFGMSVSGGLDFNGGLADITVGSRDSAVV 665
QY 597 LRSQPLRVKALMEFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
DB 666 LRSRPVVDLTYSMTETP-----DALPMVFIGKM--DVNLCFEVDSSVASEFGLREM 715
QY 654 QIQSVVTVYDLALDLSGRPHSRVFNETKNSTRQTOVLGLTQC----- 696
DB 716 FLNFTVDV-----TKORQLQCEDSSGCCSCLRKMKNGGSLCFCEHFWLI 760
QY 697 ETLKLQPLNCIEDPSPVILRLNLSLVGTPLPSAFGNLR-----PVLAEQAORLFTALP--P 750
DB 761 STEEL-----CEDCFSNITIKVYE-----FQTSGRDRDYNPTL--DHYKEPSAIFOLP 809
QY 751 FEKNGCNDNICODDLSITFSFMSLDCVLVGGPREFNVTVTNRNCGEDSVRTQVTFEPLD 810
DB 810 YEKDCKNVFCIAEIQLTN-ISQELVGVTVKVTMNLISLTSNGEDSYMTNMLNYPN 868
QY 811 LSYRKVSTLQNRQSRQSRWRLACESASSTEVSGAKSTSCSINHPIFPENSEVTFNITFDV 870
DB 869 LQFKKI-----QKPVSPDVQCDPKPV---ASVLVWCKIGHPIL-KRSSVNVSVTQQL 918
QY 871 DSKASLGNKLLKANNVTSENMPRNKTEFQLELPKYAVVYVWVTSHGVSSTKYLNTASE 930
DB 919 EESVFPNRTADITVTISNSNEKSLARETR---SLQRFHAFIAVLSR--PSVWYMN--TSQ 971
QY 931 NTSRVWQHQQVQNSLQGRSLPISLVFLVPLRNQFVIWDRPOVTFSENLSST-----CHT 985
DB 972 SPSDHKEFFNVHGENLFGAVFQLOICVPIKLQDF-----QIVRVKMLTKTDHTECTQ 1025
QY 986 KERLPSSHDSFLAELKAPVWNCISIAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTS 1045
DB 1026 SOEPACGSDPVQHVKEHWSVVCAL-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLLIVSTA-----EILFNDSVFTLLPGQAFVRSQTETKVEFF-----EVPNPLPLIV 1095
DB 1066 HTKQLLRDVSELPILGEISFNKSLYEGLNAE-----NHRTKITVIFLKEEETSLPLII 1119
QY 1096 GSSVGGLLLLALITAAVLKGLFFKQYKDMNSE 1128
DB 1120 GSSIGGLLVLVIIAILFKCGFFKRYQQLNLE 1152

RESULT 14
OB8341 PRELIMINARY; PRT; 1167 AA.
AC OB8341;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWF A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.2978 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKRQYKMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	19	Human Beta-integrin
2	5839	99.3	1153	21	Human CD11b protei
3	5839	99.3	1153	23	Human Beta2 integr
4	5839	99.3	1153	23	Human integrin 1 a
5	5839	99.3	1153	23	Integrin Mac-1 alp
6	5829	99.1	1153	11	Alpha subunit of M
7	3437	58.5	1163	11	p150.95 alpha subu
8	3423	58.2	1163	19	Human Beta-integr
9	3423	58.2	1163	21	Human CD11c protei

10	3423	58.2	1163	23	ABG61470	Human Beta2 integr
11	3421	58.2	1163	24	ABU07406	Protein differenti
12	3388	57.6	1161	16	AAW78166	Human Beta-2 integ
13	3388	57.6	1161	18	AAW23049	Human Beta 2 integ
14	3388	57.6	1161	19	AAW72825	Human alpha-d. Ho
15	3388	57.6	1161	19	AAW65089	Human Beta-integri
16	3388	57.6	1161	19	AAW57491	Human beta2 integr
17	3388	57.6	1161	20	AAW73342	Human alphad prote
18	3388	57.6	1161	21	AAW07359	Human alpha d clon
19	3388	57.6	1161	23	ABG61468	Human Beta2 integr
20	3372.5	57.4	1161	18	AAW23064	Human beta 2 integ
21	3372.5	57.4	1161	19	AAW72837	Human alpha-d deri
22	3372.5	57.4	1161	19	AAW65106	Human Beta-integri
23	3372.5	57.4	1161	20	AAW73343	Human alphad prote
24	3372.5	57.4	1161	21	AAW07376	Human alpha d prot
25	3372.5	57.4	1161	23	ABG61485	Human Beta2 integr
26	3215.5	54.7	1161	16	AAW78169	Rat alpha-d subuni
27	3213.5	54.7	1161	18	AAW23062	Rat beta 2 integr
28	3213.5	54.7	1161	19	AAW72824	Rat alpha-d #1. R
29	3213.5	54.7	1161	19	AAW60004	Rat alpha d polype
30	3213.5	54.7	1161	21	AAW07374	Rat alpha d protei
31	3213.5	54.7	1161	23	ABG61483	Rat Beta2 integrin
32	3206.5	54.5	1161	19	AAW65104	Rat beta-integrin
33	3206.5	54.5	1161	20	AAW73345	Rat alphad protein
34	3201	54.4	1161	18	AAW23061	Mouse beta 2 integ
35	3201	54.4	1161	19	AAW72836	Mouse alpha-d #2.
36	3201	54.4	1161	19	AAW65103	Mouse beta-integri
37	3201	54.4	1161	19	AAW60003	Mouse alpha d poly
38	3201	54.4	1161	20	AAW73347	Mouse alphad prote
39	3201	54.4	1161	21	AAW07373	Mouse alpha d prot
40	3201	54.4	1161	23	ABG61482	Mouse Beta2 integr
41	3197	54.4	1161	16	AAW78168	Mouse alpha-d subu
42	3192.5	54.3	1151	18	AAW23059	Rat beta 2 integr
43	3192.5	54.3	1151	19	AAW72834	Rat alpha-d #2. R
44	3192.5	54.3	1151	19	AAW65101	Rat beta-integrin
45	3192.5	54.3	1151	19	AAW60001	Rat alpha d polype

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; Protein; 1153 AA.
XX
AC AAW65090;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11b subunit protein.
XX
DE Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.

XX Homo sapiens.
XX
XX US5728533-A.
XX
XX 17-MAR-1998.
XX
XX
XX 07-JUN-1995; 95US-0485618.
XX
PR 07-JUN-1995; 95US-0485618.
PR 23-DEC-1993; 93US-0173497.
PR 05-AUG-1994; 94US-0286889.
PR 21-DEC-1994; 94US-0362652.
XX
PA (ICOS-) ICOS CORP.
XX
XX Gallatin WW, Van DER VIEREN M;
PI
XX

DR WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using

PT immobilised or labelled alpha-d polypeptide, useful for, e.g.

PT treating type-1 diabetes

XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is

CC used to describe a method for identifying compounds that modulate the

CC interaction of the beta-integrin alpha-d subunit with a binding partner

CC of alpha-d which involves contacting an alpha-d polypeptide with an

CC alpha-d binding partner, one of which is immobilised and the other of

CC which is labelled, in the presence of a test compound, and determining if

CC the compound affects binding between the alpha-d polypeptide and alpha-d

CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment

CC comprising the cytoplasmic, transmembrane or extracellular domain of

CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,

CC asthma, psoriasis, lung inflammation, acute respiratory distress

CC syndrome and rheumatoid arthritis.

XX

SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 19; Length 1153;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLTENAMTFOENARGQSVQVLOGSVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60

DB 17 FNLTENAMTFOENARGQSVQVLOGSVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMVSLGLSLAATTPPQLLACGPTVHTCSENTYVYKGLCFLFGSNLRQOPQK 120

DB 77 RLQVPEAVNMVSLGLSLAATTPPQLLACGPTVHTCSENTYVYKGLCFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLPFLMOYSEEF 180

DB 137 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLPFLMOYSEEF 196

QY 181 RIHFTFEFONNPRSLIKPITOLLGRTHPTATGLRKVVRELFINITNGARKNAFKILFL 240

DB 197 RIHFTFEFONNPRSLIKPITOLLGRTHPTATGLRKVVRELFINITNGARKNAFKILV 256

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVGLFGDAFRSEKSKQELNTVASKPRDRHVQAN 300

DB 257 TDGEKFGDPLGYEDVIPEDREGVIRYVGLFGDAFRSEKSKQELNTVASKPRDRHVQAN 316

QY 301 NFEALKTVQNLREKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLSVGSYDAG 360

DB 317 NFEALKTVQNLREKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLSVGSYDAG 376

QY 361 GVFLYTSKSTFFINMTVRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420

DB 377 GVFLYTSKSTFFINMTVRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYOHIGLVAMFR 436

QY 421 QNTGWESNANVKGTCIAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQVVCPL 480

DB 437 QNTGWESNANVKGTCIAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQVVCPL 496

QY 481 PRGQARVQCDALVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 540

DB 497 PRGQARVQCDALVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGACHVLLLRSQ 600

DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGACHVLLLRSQ 616

QY 601 PVLRYKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660

DB 617 PVLRYKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRAVFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720

DB 677 YDLALDSGRPHSRAVFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFNGNLRPVLAEDAORLFTALPPFKKNCNDNICQDDLISITFSMSLDCLVVG 780

DB 737 SLVGTPLSAFNGNLRPVLAEDAORLFTALPPFKKNCNDNICQDDLISITFSMSLDCLVVG 796

QY 781 GRPEFNVTVVRNDGEDSVRTQVTPFFPLDLVSRKVSTLQNSORSQSWRLACESASSTEV 840

DB 797 GRPEFNVTVVRNDGEDSVRTQVTPFFPLDLVSRKVSTLQNSORSQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANYVTSENNMPRTNKTEF 900

DB 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANYVTSENNMPRTNKTEF 916

QY 901 QLELPKYAVYVMTSHGKSTYKLYNFTASENRSVMOHQYOVSNLQORSLSPLSLVPLVPV 960

DB 917 QLELPKYAVYVMTSHGKSTYKLYNFTASENRSVMOHQYOVSNLQORSLSPLSLVPLVPV 976

QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNVCSIAVCORICDIP 1020

DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNVCSIAVCORICDIP 1036

QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTABILFNDSVFTLLPGQGFVRSQDET 1080

DB 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTABILFNDSVFTLLPGQGFVRSQDET 1096

QY 1081 KVEPPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1137

DB 1097 KVEPPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; Protein; 1153 AA.

XX AC AAB07360;

XX DT 17-JAN-2001 (first entry)

XX DE Human CD11b protein sequence.

XX KW Human; macrophage infiltration inhibition; alpha d integrin;

XX KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

XX KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

XX KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

XX KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

XX KW rheumatoid arthritis; central nervous system injury; CD11b.

XX OS Homo sapiens.

XX PN WO200029446-A1.

XX PD 25-MAY-2000.

XX PF 16-NOV-1999; 99WO-US27139.

XX PR 16-NOV-1998; 98US-0193043.

XX PR 08-JUL-1999; 99US-0350259.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin MW, Van Der Vieren M;

XX DR WPI; 2000-387751/33.

XX PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

XX PT macrophage infiltration and reduce inflammation at central nervous

XX PT system injury sites

XX PS Example 5; Fig 1; 270pp; English.

XX CC Integrins are a class of membrane-associated molecules that participate

in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AA60014 and XAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha d. The Alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

737	SLVGTPLSAGNLRPVLAEDAQRFLTALPFPEKQKQNDNI	IQDDLSITTFSMELDCLVWG	736
781	GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKYSTILQNQRQSRQSWRLACESASSTEV	840	
797	GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKYSTILQNQRQSRQSWRLACESASSTEV	856	
841	SGALKSTSCSINHPIIPENSESVTFNITFDVDSKASIGNKLLLKANTYSENMMPTNKTEF	900	
857	SGALKSTSCSINHPIIPENSESVTFNITFDVDSKASIGNKLLLKANTYSENMMPTNKTEF	916	
901	QLELPVKYAYVMVTVTSHGYSTKYLNFTASENTSRVMOHQYQVSNLQORSLLPISLVLVPV	960	
917	QLELPVKYAYVMVTVTSHGYSTKYLNFTASENTSRVMOHQYQVSNLQORSLLPISLVLVPV	976	
961	RLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1020	
977	RLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1036	
1021	FFGIQIEEFNATLKGNIUSFDWYIKTSHNHLLIVSTAILFNDSVFTLLPGQCAFVRQSTET	1080	
1037	FFGIQIEEFNATLKGNIUSFDWYIKTSHNHLLIVSTAILFNDSVFTLLPGQCAFVRQSTET	1096	
1081	KVEPFEPVNPPLPIVGSVGGLLLLALITAALYKLGFFFKRQYKDMMSSEGGPPGAEPQ	1137	
1097	KVEPFEPVNPPLPIVGSVGGLLLLALITAALYKLGFFFKRQYKDMMSSEGGPPGAEPQ	1153	

RESULT 3
ABG61469

ABG61469
ID ABG61469 standard: Protein: 1153 AA.

AA ABG61469;
AC

DT 27-AUG-2002 (first entry)

Human Beta2 integrin alphaCD11b subunit.

XX Beta2 integrin; alphaB subunit; CD11c subunit; CD11b subunit;
 KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.

XX Homo sapiens.

XX
PN WO200230980-A2

XX
PD
18-APR-2002

XX
PE 15-OCT-2001. 2001WQ-VIS32059XX
PP 13-0CT-2000: 2000US: 0688307XX
PA (TCOS-) TCOS CORPXX
PI Callatin WM Van Der Vieren M.

XX
WP: 2003-463260/49

XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -

XX
PS Example 5: Page 191-194. 270nn. English

XX The invention relates to promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury by administering

CC anti-alpha_d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC spinal cord injury victim. The method also involves the use of
CC a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion
CC molecule, vascular cell adhesion molecule). The method is useful for
CC promoting locomotor recovery, inhibiting locomotor damage, limiting
CC locomotor impairment, or limiting autonomic and sensory dysfunction
CC following spinal cord injury. In particular, the spinal cord injury
CC comprises compression of the spinal cord. The antibodies are also useful
CC for reducing inflammation at the site of a central nervous system injury.
CC The specification also details the identification of Beta2 integrin
CC alpha2 cDNAs and proteins, for use in raising the antibodies. Beta2
CC integrins are implicated in diseases such as LAD (leukocyte adhesion
CC deficiency, inflammatory response, diabetes, multiple sclerosis,
CC arthritis, graft atherosclerosis, inflammatory bowel disease,
CC Crohn's disease, ulcerative colitis, immune complex alveolitis
CC and leukaemia. The present sequence is a Beta2 integrin alpha subunit
CC sequence included for comparison with the Beta2 integrin alpha2 protein
CC sequences.
XX
SQ Sequence 1153 AA;
Query Match 99.3%; Score 5839; DB 23; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARFGQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMSLGLSLAATTSPPQLLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVAVNMSLGLSLAATTSPPQLLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 136
Qy 121 FPEALRGCPQSDIAFLDGGSGIIIPDPRMKEVSTIMEQLKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPQSDIAFLDGGSGIIIPDPRMKEVSTIMEQLKSKTFLSLMQYSEEF 196
Qy 181 RIHFTFEFQNNPRSLIKETITQLGRTHATGLRKVVRLEFNITNGARKNAFKILFL 240
Db 197 RIHFTFEFQNNPRSLIKETITQLGRTHATGLRKVVRLEFNITNGARKNAFKILV 256
Qy 241 TDGKFKGDPGYEDVIELDREGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
Db 257 TDGKFKGDPGYEDVIELDREGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 316
Qy 301 NFEALKTVQNLREKIFALECTGTGSSSFHEHMSQGFSAATISNGPLLSVGSYDWAG 360
Db 317 NFEALKTVQNLREKIFALECTGTGSSSFHEHMSQGFSAATISNGPLLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLGAPHYVEQTRGGQSVCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLGAPHYVEQTRGGQSVCP 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGTSPSHSQRISAGSLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGHVLLRSQ 600
Db 557 HGTSGSGTSPSHSQRISAGSLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGHVLLRSQ 616
Qy 601 PVLKVAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRREGIQSVVT 660
Db 617 PVLKVAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRREGIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRTROTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRTROTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736

Qy 721 SLVGTPLSAFGNLRPVLAEADAQRFLTALPFPFKNCGNDNICQDDLITTSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRPVLAEADAQRFLTALPFPFKNCGNDNICQDDLITTSFMSLDCLVVG 796
Qy 781 GRPEFNVTYVRNDGEDSVYRTQVTFEPLDLSYRKVSTLQNRQSRQSRWRLACESASSTEV 840
Db 797 GRPEFNVTYVRNDGEDSVYRTQVTFEPLDLSYRKVSTLQNRQSRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANTVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANTVTSENNMPTNKTEF 916
Qy 901 QLELPKYAVYVMTSHGVSSTYKLNFTASENTSRVMOHQYQVSNLQSRSLPISLVLPV 960
Db 917 QLELPKYAVYVMTSHGVSSTYKLNFTASENTSRVMOHQYQVSNLQSRSLPISLVLPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCORIOCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCORIOCDIP 1036
Qy 1021 FFGIOBEFNATLKNLSFDWYIKTSHNHLILVSTABILFNDSVFTLLPGQAFVRSQTET 1080
Db 1037 FFGIOBEFNATLKNLSFDWYIKTSHNHLILVSTABILFNDSVFTLLPGQAFVRSQTET 1096
Qy 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMMSSEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMMSSEGGPPGAEPQ 1153
RESULT 4
AAU80252
ID AAU80252 standard; Protein; 1153 AA.
XX
AC AAU80252;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human integrin 1 alpha-M subunit protein.
XX
KW Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key
FT Misc-difference 499..500
FT /note= "Encoded by GGG CAG AGG"
XX
XX WO200218583-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US27227.
XX
PR 01-SEP-2000; 2000US-229700P.
XX
PA (BLOO-) CENT BLOOD RES INC.
XX
PI Springer TA, Shimoaka M, Lu C;
XX
DR WPI; 2002-382964/41.
XX
DR N-PSDB; ABK50046.
XX
PT Modified integrin-I or integrin I-like domain polypeptide useful as an
PT immunogen to produce antibodies specific to polypeptide, comprises a
PT disulfide bond such that polypeptide is stabilized in a desired
PT conformation
XX
PS Disclosure; Page 109-112; 112pp; English.
XX
CC This invention relates to a modified integrin-I or integrin I-like

CC domain polypeptide comprising at least one disulfide bond so that the
CC domain is stabilised in a desired conformation. The polypeptide of
CC the invention may have antiinflammatory or immunosuppressive activities.
CC The polypeptides of the invention have an open conformation and are
CC useful as immunogens to produce antibodies that selectively bind to
CC integrin I-domain; and for identifying a modulator of integrin activity,
CC or of interaction of an integrin and a cognate ligand. The polypeptide
CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
CC is useful for treating or preventing an integrin mediated disorder which
CC is an inflammatory or autoimmune disorder in a subject and for
CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
CC disease, nephritis; human immunodeficiency virus (HIV), myocardial
CC infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis.
CC A therapeutic composition comprising the peptide of the invention is
CC useful for treating an integrin mediated disorder in a subject. The
CC polypeptides and/or active or antigenic fragments are useful as
CC reagents for diagnosis of integrin-mediated disorders. The present
CC sequence represents the human integrin-1 alpha-M protein subunit used to
CC generate the mutant polypeptides of the invention.
XX
SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 23; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY '1 FNLDTENAMTFQENARGFCQSVVQLQGSRRVVGAPQEIIVAAQNRGSLYOCYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFCQSVVQLQGSRRVVGAPQEIIVAAQNRGSLYOCYSTGSCPEI 76

QY 61 RLQVPVEAVNMSGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 120
DB 77 RLQVPVEAVNMSGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMQYSEEF 196

QY 181 RIHFTFKFQNNPRLSKPIITQLLGRTHATGLRKVVRELPNITNGARKNAKILFL 240
DB 197 RIHFTFKFQNNPRLSKPIITQLLGRTHATGLRKVVRELPNITNGARKNAKILVVI 256

QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVFOAN 300
DB 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVFOAN 316

QY 301 NFEALKTQVQUREKIFALEGTQTGSSSFHEMSQEGFSAITNSGPLLSTVGSVDWAG 360
DB 317 NFEALKTQVQUREKIFALEGTQTGSSSFHEMSQEGFSAITNSGPLLSTVGSVDWAG 376

QY 361 GVPLYTSKEKSTFINNTRVDSQNDNAYLCYAAAILLRNRVQSVILGAPYQHIGLVAMER 420
DB 377 GVPLYTSKEKSTFINNTRVDSQNDNAYLCYAAAILLRNRVQSVILGAPYQHIGLVAMER 436

QY 421 QNTGMWESNANVKGTQIGAFASLCSDVDVNGSDTLVLIGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTQIGAFASLCSDVDVNGSDTLVLIGAPHYEQTRGGQSVCP 496

QY 481 PRGQARWQCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYILF 540
DB 497 PRGQARWQCDVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYILF 556

QY 541 HGTSGSISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600
DB 557 HGTSGSISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 616

QY 601 PVLRVKAIINEFPREVARNVFECNDQVKGKAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
DB 617 PVLRVKAIINEFPREVARNVFECNDQVKGKAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

QY 661 YDLALDSGRPHRAVNETKNSRTRQTVLGTQTCTETIKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHRAVNETKNSRTRQTVLGTQTCTETIKLQLPNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFQNLRLPVLAEADAQRLLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFQNLRLPVLAEADAQRLLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG 796

QY 781 GPRFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
DB 797 GPRFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 856

QY 841 SGALKSTSCSNHPIFFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSNHPIFFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916

QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRSRVMOHQYQVSNLQSRSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRSRVMOHQYQVSNLQSRSLPISLVFLVPV 976

QY 961 RLNQTVIWDROVTFSENLSTCHTKERLPSHDSFLAELRKAPVNVNCSIAVCQRIQCQIP 1020
DB 977 RLNQTVIWDROVTFSENLSTCHTKERLPSHDSFLAELRKAPVNVNCSIAVCQRIQCQIP 1036

QY 1021 PFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFAVRSQTE 1080
DB 1037 PFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFAVRSQTE 1096

QY 1081 KVEPPEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPPQ 1137
DB 1097 KVEPPEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPPQ 1153

RESULT 5
AAO14428
ID AAO14428 standard; protein; 1153 AA.
XX AC AAO14428;
XX DT 03-MAY-2002 (first entry)
XX DE Integrin Mac-1 alpha subunit.
XX KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.
XX OS Unidentified.
XX PN WO200204521-A2.
XX PD 17-JAN-2002.
XX PF 09-JUL-2001; 2001WO-US21805.
XX PR 07-JUL-2000; 2000US-216600P.
XX PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX PA (BLOO-) CENT BLOOD RES.
XX PI Springer T;
XX PI WPI; 2002-148167/19.
XX PT New integrin I domain protein having alteration in at least 2
PT noncontiguous regions and exits in an open conformation, useful for
PT treating, preventing or suppressing inflammatory or immunological
PT disorders
XX PS Example 1; Fig 1F; 90pp; English.
XX CC The invention comprises structurally biased variant integrin inserted (1)
CC domain proteins, wherein the alterations to the protein occur in at least

CC two noncontinuous regions. Specifically the the variant integrin I domain
CC proteins are structurally biased to exist in the open conformation,
CC thereby altering the binding ability of the protein. The invention also
CC comprises nucleic acids encoding the variant integrin I domain proteins.
CC The integrin I domain proteins and nucleic acids are useful for treating,
CC preventing or suppressing integrin related inflammatory and immunological
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
CC proteins and nucleic acids can also be used for treating: lechaemia/
CC reperfusion (e.g. hypovolemic shock); infection; cerebral shock; viral
CC infection; and cancer. The variant integrin I domain nucleic acids and
CC proteins may be used in gene therapy, as vaccines and to screen for
CC bioactive agents. The present amino acid sequence represents the Mac-1
CC alpha subunit of integrin.

SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 23; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVWLOGSRVVGAPQEIIVAAANQSGSLVQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVWLOGSRVVGAPQEIIVAAANQSGSLVQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCTSENTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCTSENTYVKGCLFLFGSNLRQOPQK 136

Qy 121 PPEALRGCPQSDSIAFLIDGSGSIIPDPRMKELVSTIMEQLKKSTLPSLMQYSEEF 180
Db 137 PPEALRGCPQSDSIAFLIDGSGSIIPDPRMKELVSTIMEQLKKSTLPSLMQYSEEF 196

Qy 181 RIHTTFKEFQNNPRSLIKPITOLGRTHATGLRVKRVRELFNITGARKNAKFLFL 240
Db 197 RIHTTFKEFQNNPRSLIKPITOLGRTHATGLRVKRVRELFNITGARKNAKFLV 256

Qy 241 TDGKFKGDPGLGYEDVPELDREGVIRYVGLFGDAFRSEKSKQELNTVASKPPRDHVFQAN 300
Db 257 TDGKFKGDPGLGYEDVPELDREGVIRYVGLFGDAFRSEKSKQELNTVASKPPRDHVFQVN 316

Qy 301 NFEALTKVQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAATISNPGPLSTVGSYDWAG 360
Db 317 NFEALTKVQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAATISNPGPLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFTNMTVDSDMDAYLGVAALILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFTNMTVDSDMDAYLGVAALILNRVOSLVLGAPRYOHIGLVAMFR 436

Qy 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYTYEQTRGGQVSVCP 480
Db 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYTYEQTRGGQVSVCP 496

Qy 481 PRGORARQCDALVYGEQGPWGRFGAALTVLGDVNGDKLTVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARQCDALVYGEQGPWGRFGAALTVLGDVNGDKLTVAIGAPGEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMGLVDLTVGAQGHVLLRSQ 616

Qy 601 PVLVRKAIEMFNPREVARNVECNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIEMFNPREVARNVECNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSAVFNETKNSRQTQVGLGTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSAVFNETKNSRQTQVGLGTQTCETLKLQLPNCIEDPVPVILRLNF 736

Qy 721 SLVGTPLSAFGNLRVLAEDAQRFTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRVLAEDAQRFTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQNRQSRWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASIGNKLLKANTVSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASIGNKLLKANTVSENMPRTNKTFF 916
Qy 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHVQVSNLQORSLSPLSLVLPV 960
Db 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHVQVSNLQORSLSPLSLVLPV 976

Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTTERLPSPHSDFLAELRKAPVNVCSIAVCORIQCIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTTERLPSPHSDFLAELRKAPVNVCSIAVCORIQCIP 1036

Qy 1021 FFGIQEENATLKGNSLFDWYIKTSHNHLIIYSTABILFNDSVFTLLPGOGAFVRSQTET 1080
Db 1037 FFGIQEENATLKGNSLFDWYIKTSHNHLIIYSTABILFNDSVFTLLPGOGAFVRSQTET 1096

Qy 1081 KVEPPEVPNPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPO 1137
Db 1097 KVEPPEVPNPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPO 1153

RESULT 6
AAR04136
ID AAR04136 standard; protein; 1153 AA.
XX AAR04136;
AC AAR04136;
XX 25-MAR-2003 (updated)
DT 07-SEP-1990 (first entry)
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX non-specific defence system; integrin gene superfamily.
XX synthetic.

Key Modified-site 86..88 Location/Qualifiers
FT /*label= putative N-glycosylation site
FT Modified-site 240..242
FT /*label= putative N-glycosylation site
FT Modified-site 391..393
FT /*label= putative N-glycosylation site
FT Modified-site 469..471
FT /*label= putative N-glycosylation site
FT Modified-site 693..695
FT /*label= putative N-glycosylation site
FT Modified-site 697..699
FT /*label= putative N-glycosylation site
FT Modified-site 735..737
FT /*label= putative N-glycosylation site
FT Modified-site 802..804
FT /*label= putative N-glycosylation site
FT Modified-site 881..883
FT /*label= putative N-glycosylation site
FT Modified-site 901..903
FT /*label= putative N-glycosylation site
FT Modified-site 912..914
FT /*label= putative N-glycosylation site
FT Modified-site 941..943
FT /*label= putative N-glycosylation site
FT Modified-site 947..949
FT /*label= putative N-glycosylation site
FT Modified-site 979..981
FT /*label= putative N-glycosylation site
FT Modified-site 994..996
FT /*label= putative N-glycosylation site
FT Modified-site 1022..1024
FT /*label= putative N-glycosylation site

FT Modified-site 1045..1047 /*label= putative N-glycosylation site
 FT Modified-site 1051..1053 /*label= putative N-glycosylation site
 FT Modified-site 1076..1078 /*label= putative N-glycosylation site
 FT region 1..16 /*label= signal peptide
 FT region 1106..1134 /*label= putative transmembrane region
 XX BP364690-A.
 XX
 PD 25-APR-1990.
 XX
 PF 17-AUG-1989; 89EP-0115159.
 XX
 PR 23-AUG-1988; 88US-0235353.
 PR 09-MAR-1989; 89US-0321239.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Springer TA, Corbi A;
 XX
 DR WPI; 1990-125938/17.
 DR N-PSDB; AAQ04043.
 XX
 PT New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
 PT inflammation and viral infections, and in diagnosis
 XX
 PS Disclosure; Page ?; ?pp; English.
 XX
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue.
 CC Mac-1 is a member of the Integrin Gene superfamily.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1153 AA;
 Query Match 99.1%; Score 5829; DB 11; Length 1153;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1123; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
 QY 1 FNLDTENAMTFQENARFGQSVVQLOGSRVYVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 60
 Db 17 FNLDTENAMTFQENARFGQSVVQLOGSRVYVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQPOK 120
 Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQPOK 136
 QY 121 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKTLFSLMQYSEEF 180
 Db 137 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKTLFSLMQYSEEF 196
 QY 181 RIHFTKEFQNNPNRSLKPIITQLLORTHTATGLRKVYRELNIYNGARKNAFKLFL 240
 Db 197 RIHFTKEFQNNPNRSLKPIITQLLORTHTATGLRKVYRELNIYNGARKNAFKLFL 256
 QY 241 TDGEKFGDPLGYEDVPELDREGVIRVVLGFGDAFSEKSRQELNTVASKPPRDHVFQAN 300
 Db 257 TDGEKFGDPLGYEDVPELDREGVIRVVLGFGDAFSEKSRQELNTVASKPPRDHVFQAN 316
 QY 301 NFPAKTVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATITSGPLLLSTVGSYDWAG 360
 Db 317 NFPAKTVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATITSGPLLLSTVGSYDWAG 376
 QY 361 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAIAIILRNVRQSLVLGAPRYQHIGLVAMPR 420
 Db 377 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAIAIILRNVRQSLVLGAPRYQHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQSVSCPL 480
 Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQSVSCPL 496
 QY 481 PRQBARWQCDVAVLYGEOQPGWRGGAALTIVLDVNGDKLTDAIIGAPGEEDNRGAVYLF 540
 Db 497 PRQBARWQCDVAVLYGEOQPGWRGGAALTIVLDVNGDKLTDAIIGAPGEEDNRGAVYLF 556
 QY 541 HGTSGSIGSPSHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVCAGQGHVLLLRQ 600
 Db 557 HGTSGSIGSPSHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVCAGQGHVLLLRQ 616
 QY 601 PVLRYKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
 Db 617 PVLRYKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
 QY 661 YDLALDSGRPHSRVFNETKSTRQTQVLGTLTQTCETLKLQIPNCIEDPVPVILRLNF 720
 Db 677 YDLALDSGRPHSRVFNETKSTRQTQVLGTLTQTCETLKLQIPNCIEDPVPVILRLNF 736
 QY 721 SLVGTPLSAFAGNLRLPVLAEADQRLFTALFPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
 Db 737 SLVGTPLSAFAGNLRLPVLAEADQRLFTALFPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSRMELACESASSTEV 840
 Db 797 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSRMELACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 900
 Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 916
 QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASNTSRVMQHOYQVSNLQORSLPISLVFLVPV 960
 Db 917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASNTSRVMQHOYQVSNLQORSLPISLVFLVPV 976
 QY 961 RLNQTVIWDROPVTFSENLSTCTKERLPSSHDSFLAELRKAPVNVNCISIAVCQRIQCDIP 1020
 Db 977 RLNQTVIWDROPVTFSENLSTCTKERLPSSHDSFLAELRKAPVNVNCISIAVCQRIQCDIP 1036
 QY 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEIILFNDVSVFTLLPQCGAFVRSQTET 1080
 Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEIILFNDVSVFTLLPQCGAFVRSQTET 1096
 QY 1081 KVEPEVEPNPLPIVGVSSVGGLLALITAAALYKLGFFKRYKQKDMWSEGGPGASPO 1137
 Db 1097 KVEPEVEPNPLPIVGVSSVGGLLALITAAALYKLGFFKRYKQKDMWSEGGPGASPO 1153
 RESULT 7
 AAR07120
 ID AAR07120 standard; protein; 1163 AA.
 XX
 AC AAR07120;
 DT 25-MAR-2003 (updated)
 DT 05-FEB-1991 (first entry)
 XX
 DE p150.95 alpha subunit encoded by clone lambdaX47.
 KW p150,95 leucocyte adhesion receptor alpha-subunit;
 KW hairy cell leukaemia; rhinovirus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 61..63 /label=glycosylation site
 FT Modified-site 89..91 /label=glycosylation site
 FT Modified-site 385..387 /label=glycosylation site
 FT Modified-site /label=glycosylation site

OS Homo sapiens.
XX WO200029446-A1.
XX 25-MAY-2000.
XX 16-NOV-1999; 99WO-257139;
XX 16-NOV-1998; 98US-0193043;
XX 08-JUL-1999; 99US-0350259;
XX (ICOS-) ICOS CORP.
XX Gallatin MW, Van Der Vieren M;
XX WPI; 2000-387751/33.
XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
XX macrophage infiltration and reduce inflammation at central nervous
XX system injury sites
XX Example 5; Fig 1; 270pp; English.
XX Integrins are a class of membrane-associated molecules that participate
XX in cellular adhesion. Integrins are made up of an alpha subunit and a
XX beta subunit. One class of human integrins are restricted to expression
XX in white blood cells and have a common beta2 subunit; the leukocyte
XX integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
XX have an important role in immune and inflammatory responses. The present
XX protein sequence is the human integrin alpha subunit CD11c. This
XX sequence was used in an alignment to identify a novel beta2 integrin
XX alpha subunit; alpha d (AA060014 and AAB07359). The present sequence has
XX approximately 66% identity to the protein sequence of alpha d. The
XX Alpha_d gene and protein may be useful in therapy for diseases linked
XX to aberrant alpha d function e.g. Type I diabetes, atherosclerosis,
XX multiple sclerosis, asthma, psoriasis, lung inflammation, acute
XX respiratory distress syndrome, rheumatoid arthritis and leukocyte
XX adhesion deficiency (LAD). In addition, anti-alpha d monoclonal
XX antibodies may be used in the inhibition of macrophage infiltration at
XX the site of a central nervous system injury. The monoclonal antibodies
XX can also be used to detect and diagnose Crohn's disease.
XX Sequence 1163 AA;
SQ
Query Match 58.2%; Score 3423; DB 21; Length 1163;
Best Local Similarity 60.2%; Pred. No. 3.2e-274;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFOENARGFQGSVVVQLQGSRRVVVGAPOEIVAAHQGSLYQCDYSTGSCPEI 60
DB 20 FNLDTEELTAFRVDSAGRGDSVVQVANSWVVGAFQKIIAANOIGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFGLFSGNLRLQPOK 120
DB 80 GLQVPPEAVNMSLGLSLASTTSFQSLLACGPTVHCEGRNMYLTGLCFLLGPT--QLTQR 137
QY 121 FEARLGCQPOESDTAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLESLMOYSBEF 180
DB 138 LSVRSQECQPOEQDIFVLIDGSGSISSRNFAFMNFVRAVISQFORPSTQFSLMQFSNKF 197
QY 181 RIHTFKFQNNPNRSLKPTITOLLGRTHATGLRKVVRELFTNITNGARKNAKIFLL 240
DB 198 QHTFTFEPRRTSNPLSLASVHQLQGYTTATQNVVHRFLPHASYGARRDAIKILIVI 257
QY 241 TDGEKFGDPLGVEDVPELDREGVIRYVLFGDAFRSEKSRQELNTVASKPPDRHVQAN 300
DB 258 TDGKKGESLDYKDVIPMAADAAGIIRYAIGVGLAFQNRNSWELNDIASKPSQEHIFKVE 317
QY 301 NFEALKTQVQNRKFIKFAIEGTQGTSSSSFEHMSQEGSAITNSGPNLLSTVGSYDAG 360
DB 318 DFDALKDIQNLKKEKIFAEIGTETISSSSFELEMAQEGFSVFTPDGPVLGAVGFTWSG 377
QY 361 GWFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVILGAPRYQHIGLVAMFR 420

DB 378 GAFLYPPNMSPTFINNMQENVDNRDSTYLGSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437
QY 421 QNTGMENANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVSCPL 480
DB 438 QVSRQWRKAEVIGTOIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEOTRGQSVSCPL 497
QY 481 PRQARAWOCDAVLYGEGQGPWGRFGAALTGLGVNVDKLTVDVAIGAPGEEDNRGAYVLF 540
DB 498 PRGWR-RWWDCAVLYGEGQHPWGRFGAALTGLGVNVDKLTVDVIGAPGEENRGAYVLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTGVAGHVLRLRSQ 600
DB 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTGVAGHVLRLTR 616
QY 601 PVLVRKAIMEFNPREVARNVFCNDQVWKGKEAGEVRVCLVHVKSTRDRREGQIQSVVT 660
DB 617 PVLWGVSMQFIPAEIPRSFAFECREQVSEOTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
DB 677 LDLALAPGRLS PRAIFQETKNRSLRVRLGLKACENFNLLPSCVEDSVIPIILRLNF 736
QY 721 SLVGTPLSAGNLRPLVLAEDAQRLFTALFFPKXCGNDNICODDLSITFFMSLDCLVWG 780
DB 737 TLVGKPLLAFLNRLPMLAALAQRYFTASLPFKXCGADHICODNLGHSFSPGLKSLVG 796
QY 781 GPREFNVTVVRNDGEDSYRTVTFPPPLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
DB 797 SNLELNAEVMWNDGEDSYGTTTFSPAGLSRYVAEGQKQGLRSLHLTC--CSAPVG 854
QY 841 SGALKSTSCSINPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTETEF 900
DB 855 SQGTWSTFSCRINHLPFGGAQITFLATFVSPRAVGLDRLILLIANVSSENIPRTSKTIF 914
QY 901 QLELPVKYAVMYVTSHGVSTKYLNFAS--ENTSRVMQHOYOVSNLQORSIPISLFLVP 959
DB 915 QLELPVKYAVYIVVSSHEQFTKYLNFSSEKESHVAMHYQVNNLQORLQPLVSNFWVP 974
QY 960 VRLNQTWDRPOVTFSENLSSTCHTRELPSHSDFLAELRKAPVNCISAVCQRIQCDI 1019
DB 975 VELNQEAVMDVEVSHQPNFSLRCSSEKIAPPASDFLAHTQKNPVLDCSAGCLRFRCDV 1034
QY 1020 PFFGIOBEFNATLGNLSFDWYIKTSHNHLILVSTAEILFNDSVFTLLPGGAFVRSQTE 1079
DB 1035 PSFSVQEEELDFTLKGNLSFGWVRQILQKVSVSVASBIIPTSVYSQLPQGEAFPMRAQTI 1094
QY 1080 TKVPEFPEVPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMHSE 1128
DB 1095 TVLEKYKVHNPILVIGSSIGGLLLALITAVLYKVGFKKRYKEMMEE 1143
RESULT 10
ABG61470
ID ABG61470 standard; Protein; 1163 AA.
XX AC ABG61470;
XX XX
XX 27-AUG-2002 (first entry)
XX Human Beta2 integrin alphaCD11c subunit.
XX DE Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
XX KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
XX KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
XX KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
XX KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
XX KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
XX KW locomotor recovery; locomotor damage; locomotor impairment;
XX KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX OS Homo sapiens.
XX XX

PN WO200230980-A2.
XX 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32059.
XX
PR 13-OCT-2000; 2000US-0688307.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Van Der Vieren M;
PI
XX WPI; 2002-463260/49.
DR
XX
XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -
XX
PS Example 5; Page 194-198; 270pp; English.
XX
CC The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an
CC anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC spinal cord injury victim. The method also involves the use of
CC a ligand selected from ICAM-3 or VCAM-1 (intracellular cell adhesion
CC molecule, vascular cell adhesion molecule). The method is useful for
CC promoting locomotor recovery, inhibiting locomotor damage, limiting
CC locomotor impairment, or limiting autonomic and sensory dysfunction
CC following spinal cord injury. In particular, the spinal cord injury
CC comprises compression of the spinal cord. The antibodies are also useful
CC for reducing inflammation at the site of a central nervous system injury.
CC The specification also details the identification of Beta2 integrin
CC alpha d cDNAs and proteins, for use in raising the antibodies. Beta2
CC integrins are implicated in diseases such as LAD (leukocyte adhesion
CC deficiency, inflammatory response, diabetes, multiple sclerosis,
CC arthritis, graft atherosclerosis, inflammatory bowel disease,
CC Crohn's disease, ulcerative colitis, immune complex alveolitis
CC and leukaemia. The present sequence is a Beta2 integrin alpha subunit
CC sequence included for comparison with the Beta2 integrin alpha d protein
CC sequences.
XX
SQ Sequence 1163 AA;
Query Match 58.2%; Score 3423; DB 23; Length 1163;
Best Local Similarity 60.2%; Pred. No. 3.2e-274;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFQENARGFGQSVVLOGSRVVVGAPOEIVAAQNRGLYOCYSTGCEPI 60
DB 20 FNLDTEELTAFRVDSAGFGDSVVQVANSVVVVGAPQKIIPAAQIGGLYOCYSTGACEPI 79
QY 61 RLQVPVEAVNMSGLSLAATTSPPQLACGPTVHTCSENTRYVKGFLGFLGSLNRQQPQK 120
DB 80 GLQVPEAVNMSGLSLASTTSPSOLLACGPTVHCEGGRMYLTGLCFLGPT--QLTOR 137
QY 121 FPEARLGCQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMQYSEEF 180
DB 138 LPVSRQECPRQEQDVI VFLIDGSGSISSRNFAFMNFRAVISQFORPSTQFSLMQFSNKF 197
QY 181 RIHFTFKFQNNPNPSLKIPTOLLGRTHATGLRKVVRELFNITNGARKNAFLIFLL 240
DB 198 QTHFTFEEPRRTSNPLSLASVHQLOGFTYTATQNVVHRFLPHASYGARRDAIKILIVI 257
QY 241 TDGEXFGDPLGYEDVPIPEDRGVIRYVLGFGDAFSEKSRQELNTVASKPRDHFVQAN 300
DB 258 TDGKKGDSLDYKDVIPMDADAAGIRYAGVGLAQNRNSWKELDIASKPQOEHIKVE 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTSSSSFEHEMGEQGSAAITSNGLPSTVSGYDNAG 360
DB 318 DFDALKDIQNLKEKIFAIEGTETITSSSFELEMAQEGFSAVFTPDGPVLGAVGFTWSG 377

QY 361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSYTELALWKGVSQSLVGLGAPRYQHIGKAVIFI 437
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNSTDVLIGAPHYYEQTGGQSVQVCP 480
DB 438 QVSRQWRMKAEBVIGTIGSYFGASLCSVDVDTGSLDVLIGAPHYYEQTGGQSVQVCP 497
QY 481 PRGORARWQCDVLYGEOGPWGRFGAALTVDGVNVDGKLTDAIAGPGEEDNRGAVYLF 540
DB 498 PRGWR-RWVCDVLYGEOGHPWGRFGAALTVDGVNVDGKLTDAIAGPGEEDNRGAVYLF 556
QY 541 HGTSGSISPSHSQRIAGSKLSPLQYFGQSLGSGQDLTMDGLVDLTVGAQSHVLLRSQ 600
DB 557 HGVLGPSISPSHSQRIAGSKLSPLQYFGQSLGSGQDLTMDGLVDLTVGAQSHVLLRSQ 616
QY 601 PVLRYKATMEFNPREVARNFECNDQVVGKAESEVRVCLHVQKSTRDLRLEGQIQSVVT 660
DB 617 PVLWVGSMQFIPAEIPRSFAFECEQVVSQTLVQSNICLYDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRTOVLGTTCTETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 LDALAPGLRSLPRAIFQETKRSRVRVGLGKAHCENFNLLPSCVEDSVIPIILRNF 736
QY 721 SLVGTPLSAFNGNLRPVLAEADQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 TLVCKPLLAFLNRLPMLAALAAQRYFTASLPFEKNCADHICQDNLGISFSFGLKSLLVG 796
QY 781 GPRFNVTVVRNDEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV 840
DB 797 SNLELNAEVMVWMDGDSYGTITFSHPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854
QY 841 SGALKSTSCSINHPIFPENSEVTENITPDVDSKASLGNKLLKANVTSENMPRTKTEF 900
DB 855 SGTWSTSCRINHPIFRGAQITFLATPDVSKAVGLDRLLLIANVSENNIPRTSKTIP 914
QY 901 QLELPVKYAVVMVTSHTGVSTKYNLFTAS-ENTSRVMOHQYQVSNLGRSPISLVFLVP 959
DB 915 QLELPVKYAVVIVVSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSNFVWP 974
QY 960 VRLNQTVIWRDPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQCDI 1019
DB 975 VELNQEAVMMDVEVSHPNQPSLRCSSEKXIAPASDFLAHIQKNPVLDCSIAGLRFRCDV 1034
QY 1020 PFGIOEFENATLKNLSFDMVYIKTSHNHLIIVSTAELFNDLSVFTLLPGOGAFVRSOTE 1079
DB 1035 PSFSVQEBLDFTLKGNLSFGWVRQILQKVSVWSVAEIIFDTSVYSQLPGOAFMRAQTI 1094
QY 1080 TKVPEFVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKQKMMSE 1128
DB 1095 TVLEKYKVNPIPLIVGSSIGGLLLALITAVLYKVGFQYKEMMEE 1143
RESULT 11
ABU07406
ID ABU07406 standard; Protein; 1163 AA.
XX ABU07406;
XX
XX 28-JAN-2003 (first entry)
XX
XX Protein differentially regulated in prostate cancer #9.
XX
XX Prostate cancer; gene expression; differential regulation;
KW molecular marker; drug target; cancer detection; cancer diagnosis;
KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
XX Homo sapiens.
XX
XX WO200281638-A2.
XX
XX 17-OCT-2002.
XX

Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis; inflammatory bowel disease; asthma.

Homo sapiens.

Key Location/Qualifiers

Domain 17..1108 /note= "extracellular domain"

Region 150..352 /note= "this region is homologous to the insert common to CH1a,b,c and may be a site for interaction with ICAM family proteins"

Binding-site 465..474 /note= "putative cation binding site"

Binding-site 518..527 /note= "putative cation binding site"

Binding-site 592..600 /note= "putative cation binding site"

Region 1109..1128 /note= "transmembrane region"

Domain 1129..1161 /note= "cytoplasmic domain"

WO9517412-A1.

29-JUN-1995.

21-DEC-1994; 94WO-US14832.

05-AUG-1994; 94US-0286889.

23-DEC-1993; 93US-0173497.

(ICOS-) ICOS CORP.

Gallatin WM, Van Der Vieren M;

WPI; 1995-240603/31.

N-PSDB; AA091712.

Alpha sub-unit polypeptide of human beta 2 integrin - used to identify potential antiinflammatory agents, for the treatment of graft arteriosclerosis, inflammatory bowel disease, asthma, etc.

Claim 7; Page 82-87; 172pp; English.

A probe based on a partial cDNA clone (given in AA091727) of canine alpha-Tw1 was used to screen a human spleen cDNA library to identify clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d.

The cDNA was manipulated to allow expression of recombinant alpha-d subunit in COS and CHO cells.

Sequence 1161 AA;

Query Match 57.6%; Score 3388; DB 16; Length 1161;

Best Local Similarity 58.9%; Pred. No. 2.6e-271;

Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

1 FNLDTENAMTFQENARGFGQSVVLOGSRVVGCAPOEIVAAVNAORGSLYOCDYSTGSCPEI 60

17 FNLDVEEPTIFQEDAGFGQSVVQFGGSLVWVGAPLEVAANQTGRLYDCAATGMCQPI 76

61 RLQVPVEAVNMSLGLAATTSPOLLACGPTVHOTCSENTYVKGCLFLFGNLRQPOK 120

77 PLHIREAVNMSLGLTAASTNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-EIQT 135

121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKKSKTLFSLMOYSEEF 180

136 VPDATPECPHOEMDIIVFLIDGSGSIDQNDFNQMKGFQAVMGQFECTDLFALMOYSNLL 195

181 RIHFTKEQNPNPSLKPITQLLGRTHATGLRKVVRELPNITNGARKNAFKLLFIL 240

196 KIHFTTQRTSPSQSLVDPIVLQKGLTFTATGILTVVTQLPHHKNKGARKSAKILLVI 255

241 TDGEKFGDPLGYEDVIPELDREGVIRYVGLFGDAPFRSEKSRQELNTVASKPRDHVFOAN 300
 256 TDGQKYDPLEYSDVIPQERAGIIRYAIYGVGHAPQGTARQELNTISSAPPQDHVFKVD 315
 301 NFEALKTVONOLREKIFAIEGTQTGSSSSPFHEMSQEGFSAATNSNGPLLTSTVGSYDAG 360
 316 NFAALGSIQKQLEKQIYAVEGTQSRASSSFQHEMSQEGFSTALTMDDFLGAVGSFWSG 375
 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
 376 GAFLYPPNMSPTFINMSQENVMDRDSYLCYSTELALWKGVNVLGAPRYQHTGKAVIFT 435
 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYQTRGGQVSVCP 480
 436 QVSRQWRKKAQEVGTQIGSYFGASLCSVDVDSNGSTDVLVIGAPHYYQTRGGQVSVCP 495
 481 PRGQBARWOCDAVLYGEOGPWGRGAALTVLGVNGDKLTDVAIGARGEEDNRCAGVYLF 540
 496 PRGQVQWOCDAVLRGEQGPWGRGAALTVLGVNGDKLTDVAIGARGEEDNRCAGVYLF 555
 541 HGTSGGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDTLVGAQCHVLLRSQ 600
 556 HGASEGISPSHSQRIASSQLSPRLQYFGQALSGQDLTQGLMDLAVGARGQVLLRSL 615
 601 PVLRYKAIEMFNPREVARNVFECDNDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
 616 PVLKVGAMRFPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSDLQ--GDIQSSVR 673
 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 720
 674 FDLALDPGLTSLRAIFNETKPTLTKRRLTGHLGHCETLKLPLPCDVEDVVSPIILHNF 733
 721 SLVGTPLSAFGNLRPLVAEDAQRFTALFPPEKCNKNDNIQDDLSITFSFMSLDCLVVG 780
 734 SLVREPIFSPQNLRPVLAVGSDQLFTASLPPEKNCQGLCEGLGVTLFSGLQTLTVG 793
 781 GPREFNVTVVRNDCGDSYRTQVTFPFPLDLISYRKVSTLQNGRSQRSMRLACESASSTEV 840
 794 SSLELNVIIVTNAGDSYGTIVSVLYYPAGLSHRRVSGAOKPHQSALRLACETV-PTED 852
 841 SGALKSTGCSNHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
 853 EG-LRSSRCVNHPIFHEGSGNGTFTVTFDVSYKATLGDRLMLRASSENKASSSKATF 911
 901 OLELDPKAVVMVMTSHGVSTKYLNLF-TASENTSVMOHVOVSNLQORSLSPLSLVFLVP 959
 912 QLELDPKAVVMVMTSHGVSTKYLNLF-TASENTSVMOHVOVSNLQORSLSPLSLVFLVP 971
 960 VRLNQTIVWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDI 1019
 972 VLLNGVAVVDVMEAPSQSL--PCVSEKRPQHSDFLTQISRSPLMDCSIADCLQRCOV 1029
 1020 PFFGQBEFNATLKNLSFDWYIKTSHNHLIVSTAELFNDVSFTLLPGQAFVRSQTE 1079
 1030 PSFSVQBEDELFTLKNLSFGVWRETQKKLVVSVAEITFDTSVYSQLPQGEAFMRAQME 1089
 1080 TKVEPFEVNPPLPLTVGSSVGLLALLITALYKLGFFKQYKQDMWSE 1128
 1090 MVLEDEVEYNAIPIIMSGSSVALLLALLITALYKLGFFKQYKQDMWSE 1138

RESULT 13
 AAW23049
 ID AAW23049 standard; Protein; 1161 AA.
 XX AAW23049;
 AC AAW23049;
 XX AAW23049;
 DT 24-FEB-1998 (first entry)
 XX Human beta 2 integrin alpha d subunit.
 DE Human beta 2 integrin alpha d subunit; human; cell migration;
 XX Beta 2 integrin alpha d subunit; human; cell migration;
 KW cell adhesion; phagocytosis; diabetes; atherosclerosis;

multiple sclerosis; asthma; psoriasis; lung inflammation; acute respiratory distress syndrome; rheumatoid arthritis.

Homo sapiens.

[illegible]

Domain	17..1108	/label= Extracellular domain
--------	----------	------------------------------

```
/label= Extracellular domain
```

Domain	
1109..1128	

```
| /label= Transmembrane_domain
```

/note= "homologous to the human CD11c transmembrane

Domain	region"
1129	1161

DOMAIN
1129..1161
/label=Cytosplasmic domain

Domain
150..352
/ rader- cytoplasmic domain

domain: 100.0000
/note= "region homologous to the I (insertion)
domain common to CD11a, CD11b and CD11c"

W09731099-A1

29-AUG-1997

24-SEP-1997 0740-11503713

22 FEB 1967 06110 0605073

1
1
0
0
0
0
1
-
0
0
0
.,

—

—

WPI; 1997-435154/

Hybridoma 199M and antibody secreted by it - specific for new rat beta2 integrin subunit, useful to detect subunit in cells and modulate its activity

Example 5; Page 116-120; 222pp; English.

This polypeptide comprises a novel human beta 2 integrin subunit, designated alpha d. Its sequence was deduced from a cDNA clone (see AAT7920) isolated from a spleen cDNA library. Alpha d is involved in cell migration, phagocytosis and cell-cell interaction. Recombinant alpha d polypeptides can be expressed in transformed host cells for use in assays for identifying antibodies or other compounds that modulate alpha d activity or which modulate the interaction between alpha d and a ligand, for treating or preventing diseases in which macrophages are implicated. Treatment is applicable to disease states in which alpha d binding, or localised accumulation of cells which express alpha d, is implicated such as such as type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome and rheumatoid arthritis.

Sequence 1161 AA:

57 64. score 3388. DB 18: Length 1161:

Very Match	37.6%	Score 3388;	Length 1161;
1st Local Similarity	58.9% <td>Pred. No. 2.6e-271; <td></td> </td>	Pred. No. 2.6e-271; <td></td>	

Seq. local similarity	Seq. no.	Indels	Gaps
665; Conservative	288; Mismatches	288; Indels	8; Gaps
665; Conservative	168; Mismatches	288; Indels	8; Gaps

1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVWGAPQEIVAANQSGSLYQCDYSTGSCEPI 60

1	..
2	..
3	..
4	..
5	..
6	..
7	..
8	..
9	..
10	..
11	..
12	..
13	..
14	..
15	..
16	..
17	..
18	..
19	..
20	..
21	..
22	..
23	..
24	..
25	..
26	..
27	..
28	..
29	..
30	..
31	..
32	..
33	..
34	..
35	..
36	..
37	..
38	..
39	..
40	..
41	..
42	..
43	..
44	..
45	..
46	..
47	..
48	..
49	..
50	..
51	..
52	..
53	..
54	..
55	..
56	..
57	..
58	..
59	..
60	..
61	..
62	..
63	..
64	..
65	..
66	..
67	..
68	..
69	..
70	..
71	..
72	..
73	..
74	..
75	..
76	..
77	..
78	..
79	..
80	..
81	..
82	..
83	..
84	..
85	..
86	..
87	..
88	..
89	..
90	..
91	..
92	..
93	..
94	..
95	..
96	..
97	..
98	..
99	..
100	..

0 1 2 3 4 5 6 7 8 9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

61 RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHQTCSENTYVKGLCFLEFGSNLRQQPQK 120

77 PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

[illegible]

XX

KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
 KW immunoglobulin; fusion protein; binding molecule; antibody;
 KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
 KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
 KW acute respiratory distress syndrome.

XX

OS Homo sapiens.

XX

PN US581029-A.

XX

PD 03-NOV-1998.

XX

PF 07-JUN-1995; 95US-0482293.

XX

PR 07-JUN-1995; 95US-0482293.

XX

PR 23-DEC-1993; 93US-0173497.

XX

PR 05-AUG-1994; 94US-0286889.

XX

PR 21-DEC-1994; 94US-0362652.

XX

PA (ICOS-) ICOS CORP.

XX

PI Gallatin WM, Van Der Vieren M;

XX

DR WPI; 1998-609318/51.

XX

DR N-PSDB; AAV67281.

XX

PT Antibodies specific for beta-2 integrin alpha-subunit d - useful in
 PT immunohistochemical analysis

XX

PS Example 5; Column 61-66; 106pp; English.

XX

CC The present sequence represents human alpha d. The present invention
 CC describes: (1) an antibody that specifically binds alpha d, which
 CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
 CC antibody as in (1); (3) an anti-idiotype antibody specific for the
 CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
 CC used in immunohistochemical analysis to localise alpha d to subcellular
 CC compartments or individual cells within tissues. Substances that modulate
 CC alpha d binding (which may include antibodies) can be used to treat
 CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
 CC inflammation, acute respiratory distress syndrome or arthritis.

XX

SQ Sequence 1161 AA;

XX

Query Match 57.6%; Score 3388; DB 19; Length 1161;

XX

Best Local Similarity 58.9%; Pred. No. 2.6e-271;

XX

Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

QY

1 FNLDTENAMTFQENARGFGQSVVLOGSRVWVGAPQEIIVAANORGSLYOCXYSTGSCPI 60

DB

17 FNLDVEEPTIFQEDAGFGQSVVQFGSLRVVGAPELVAAQTGLYDCAATGMCQPI 76

QY

61 RLQVPVEAVNMISGLSLAATTPSPOLLACGPTVHOTCSENTYVKGCLFLFGNLROQPK 120

DB

77 PLHIRPEAVNMISGLTLAATNGSRLACGPTLHRVCGENSYSKSGCLLGSRW-ELIQT 135

QY

121 FEALRGCCQEDSDIAFLIDGSGSIIPHDFRNMKELVSTIMEQLKSKTLFSLMQVSEEF 180

DB

136 VPDATPECPHQEMDVIPLIDGSGSIDQNDNFQMKFVQAVMGQFEGDTLFFALMQVSNLL 195

QY

181 RIHFTKEFQNNPNPSLKPITOLLGRTHTATGLRKVVRELPNINGARKNAFKTLFL 240

DB

196 KIHFTIQTPTSPSQOSLVDPVOLKGLTFTATGILTVVTFPHHNGARKSAKTLI 255

QY

241 TDGEKFGDPLGEDVPIPEIDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHFVQAN 300

DB

256 TDGQKYKDPLEYSVDVPOREKAGIIRYAGVGHAFQPTARQELNTISSAPQDHFVKVD 315

QY

301 NFEALKTQVQUREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLSLVGVSDWAG 360

DB

316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLGAVGSPWSG 375

QY

361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAAIILNRNRVQSLVLCGAPRYQHIGLVAMPR 420

DB

376 GAFLPPNMPSTFINNSQENVDMDRSLGYSTELALMKGVQNLVLGAPRYQHTKAVIPT 435

QY

421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQVSVCP 480

DB

436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSGSDTLILIGAPHYYEOTRGQVSVCP 495

QY

481 PRGORARWOCDAVLVGEQOPMGRFGAALTIVLDVNGDKLTDVAIGAGEEDNRCNAVLY 540

DB

496 PRGQRVQWOCDAVLVGEQOPMGRFGAALTIVLDVNGDKLTDVAIGAGEQENRCNAVLY 555

QY

541 HGTSGSIGSPSHSQRISAGSKLSPRIQYFQCSLGGQDLTMDGLVLTVAQAGHVLRLSQ 600

DB

556 HGASESGISPSHSQRISAGSKLSPRIQYFQCSLGGQDLTMDGLMDLAVGARGQVLLRLSL 615

QY

601 PVLRVKATMEFNPREVANRVFECNDQVVKGKAGEVVRVCLVHKVQKSTRRLRREGQIOSVVT 660

DB

616 PVLKVGAMRFSPVEVAKAVRCWEKPSALEAGDATVCLTTQKSLDQL--GDIQSSVR 673

QY

661 YDLALDSGRPHSRVFNETKSTRQTVLGLTCTETLKLQLPNCIEDPPSPVILRLNF 720

DB

674 FDLALDPGLTSRAIFNETKNTLTKRKTGLGHIHCETLKLPLPCDCEVDVSPVILHLNF 733

QY

721 SLVGTPLSAFGNLRPLVLAEDAQRLETFALFPFPEKNGCNDNICODDLSITFSFMSLDCLVVG 780

DB

734 SLVREPIPSQNLRPVLA VGSQDLFTASLPFEKNGQDGLCEGDLGVTLSFGLQTLTVG 793

QY

781 GPREFNVTVRNDEGDSYRTQVTPFFPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840

DB

794 SSLELNVITVWNAGEDSYGTVSVLYPAGLSHRRVSGAQKPHQSALRLACETV-PTED 852

QY

841 SCALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900

DB

853 EG-LRSSCSNVNHPHFEGSNGTFTVTFDVSYKATLGDRLMLRASSENKASSKATF 911

QY

901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHYQVSNLQSRSLPISLVFLVP 959

DB

912 QLELPVKYAVYTMISROBESTKYFNPAFSDKKKEAEHRYRVNLSQRDLAISINFWVP 971

QY

960 VRLNQTIVWRDQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDI 1019

DB

972 VLLNGVAVMDVVMVEAPSQSL--PCVSEKRPQHSDFLTQISRSPLDCSIADCLQPRCDV 1029

QY

1020 PFGFQOEENATLKNLSFDVYIKTSHNHLIVSTAELFNDSVETLLPGQCAFVRSOTE 1079

DB

1030 PFSVQOEELDTLKNLSFGWVRRETLQKKVLVVAEITFTDTSVYSQVPGQAFMRAQME 1089

QY

1080 TKVEPEFEPNPLPIVGSVGGLLLLALITAAALYKLGFFKRYKQWMMSE 1128

DB

1090 WYLEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRYKHEMLED 1138

XX

RESULT 15

XX

AAW65089

XX

ID AAW65089 standard; Protein; 1161 AA.

XX

AC AAW65089;

XX

DT 28-SEP-1998 (first entry)

XX

Human Beta-integrin alpha-d subunit protein isolated from clone 19A2.

XX

Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;

XX

type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;

XX

lung inflammation; acute respiratory distress syndrome;

XX

rheumatoid arthritis.

OS

Homo sapiens.

XX

Key Location/Qualifiers

FH

Peptide 1..16

FT		/label= signal	
FT	Protein	17..1161	
FT		/notes="Beta-integrin alpha-D subunit"	
FT	Domain	17..1108	
FT		/label= extracellular	
FT	Domain	150..352	
FT		/notes="homologous to insertion domain of CD11 subunits"	
FT	Domain	1129..1161	
FT		/label= cytoplasmic	
XX	US5728533-A.		
PX			
XX	17-MAR-1998.		
PD			
XX	07-JUN-1995;	95US-0485618.	
XX			
XX	07-JUN-1995;	95US-0485618.	
PR	23-DEC-1993;	93US-0173497.	
PR	05-AUG-1994;	94US-0286889.	
PR	21-DEC-1994;	94US-0362652.	
XX	(ICOS-) ICOS CORP.		
PA			
XX			
PI	Gallatin WM, Van DER VIEREN M;		
XX			
DR	WPI; 1998-206565/18.		
DR	N-PSDB; AAV35236.		
XX			
PT	Screening assay for modulators of integrin binding - using		
PT	immobilised or labelled alpha-d polypeptide, useful for, e.g.		
PT	treating type-I diabetes		
XX			
XX	Example 5; Fig 1A-D; 106pp; English.		
CC	This sequence represents a novel human beta-integrin alpha-d subunit.		
CC	This sequence is used in a method for identifying compounds that modulate		
CC	the interaction of alpha-d with a binding partner of alpha-d which		
CC	involves contacting an alpha-d polypeptide with an alpha-d binding		
CC	partner, one of which is immobilised and the other of which is labelled,		
CC	in the presence of a test compound, and determining if the compound		
CC	affects binding between the alpha-d polypeptide and alpha-d binding		
CC	partner, where the alpha-d polypeptide is alpha-d or its fragment		
CC	comprising the cytoplasmic, transmembrane or extracellular domain of		
CC	alpha-d. Compounds that modulate alpha-d binding could be used to treat		
CC	diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,		
CC	asthma, psoriasis, lung inflammation, acute respiratory distress		
CC	sndrome and rheumatoid arthritis.		
XX			
SQ	Sequence 1161 AA;		
	Query Match 57.6%; Score 3388; DB 19; Length 1161;		
	Best Local Similarity 58.9%; Pred. No. 2.6e-271;		
	Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;		
Oy	1 FNLDTENAMTFQENARGFGQSIVQLQGSRVVVGAPQEIVAANQRGSLYQCDDYSTGSCEPI 60		
Dd	:		
Dd	17 FNLVDVEEPTIFEODAGGGFQSVVQFGSRLVVGAPLEVAANTGRLYDCAATGMCPPI 76		
Oy	61 RLQVPVEAVNWSLGLSLAAATSPPLLACGTVTHTCSENTYVKGLCFELFGSNLRQQPOK 120		
Dd	:		
Dd	77 PLHIRPEAVNNWSLGLTLAASNGSRLLACGTPLHRVCENSYSXGSCLLLSRW-ETIQT 135		
Oy	121 FPEARLGCGPQSDIDIAFLIDSGSSIIPHDFRMRKELSTIMEOLKKSKTFLFSMQYESEF 180		
Dd	:		
Dd	136 VPDATPECPCHEMDIVFLIDSGSSIDQNDFNQMGKFVOAVNGQEGTDITLPALM QYSNLL 195		
Oy	181 RIHFTEFKFEQNNPNRSIKETOLLGRTHRTATGLRKVKVRRELFININGARKNAFKILFL 240		
Dd	:		
Dd	196 KIHFTFTFRTPSQQSUVDPVLQKLGTFTATGITLVTVQLFHHONGARKSAKKILIVI 255		
Oy	241 TDGEKFGDPLYGYEDVIPELDREGVIKYVLGVGDGFADRFSEKSRQELNTVASKEPRDHUFOAN 300		
Dd	:		
Dd	256 TDKQYKDKPLSYSDVIPQAEEKAGIRYAIGVGHAFQGPARTQELNLTISSAPPQDFVKVD 315		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:20:45 ; Search time 26.1265 Seconds
(without alignments)
8026.803 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFCQ.....FKRQKDMSEGGFCAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5879	100.0	1137	11	US-09-902-481A-3
2	5857	99.6	1137	11	US-09-902-481A-4
3	5839	99.3	1153	9	US-09-350-259-3
4	5839	99.3	1153	11	US-09-902-481A-1
5	5839	99.3	1153	11	US-09-891-943-3
6	5839	99.3	1153	15	US-10-144-259-30
7	5839	99.3	1153	15	US-10-207-655-176
8	5836	99.3	1137	11	US-09-902-481A-5
9	5832	99.2	1137	11	US-09-902-481A-6
10	5823.5	99.1	1152	10	US-09-945-265-4
11	3446	58.6	1163	12	US-10-116-275-204
12	3423	58.2	1163	9	US-09-350-259-4
13	3423	58.2	1163	11	US-09-891-943-4
14	3388	57.6	1161	9	US-09-350-259-2
15	3388	57.6	1161	11	US-09-891-943-2

16	3372.5	57.4	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3372.5	57.4	1161	11	US-09-891-943-99	Sequence 99, Appl
18	3213.5	54.7	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3213.5	54.7	1161	11	US-09-891-943-55	Sequence 55, Appl
20	3201	54.4	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3192.5	54.3	1161	11	US-09-891-943-53	Sequence 53, Appl
22	3192.5	54.3	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3192.5	54.3	1151	11	US-09-891-943-37	Sequence 37, Appl
24	3180	54.1	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3180	54.1	1155	11	US-09-891-943-46	Sequence 46, Appl
26	1536.5	26.1	1170	10	US-09-945-265-2	Sequence 2, Appl
27	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, Appl
28	1229.5	20.9	494	11	US-09-891-943-103	Sequence 103, Appl
29	1128.5	19.2	413	9	US-09-350-259-101	Sequence 101, Appl
30	1128.5	19.2	413	11	US-09-891-943-101	Sequence 101, Appl
31	1128	19.2	1179	12	US-10-177-550-2	Sequence 2, Appl
32	1085.5	18.5	1151	11	US-09-984-130-103	Sequence 103, Appl
33	1085.5	18.5	1151	12	US-09-836-353A-103	Sequence 103, Appl
34	1079.5	18.4	1189	11	US-09-984-130-35	Sequence 35, Appl
35	1079.5	18.4	1189	12	US-09-836-353A-35	Sequence 35, Appl
36	1054	17.9	1181	15	US-10-160-354-2	Sequence 2, Appl
37	1004	17.1	216	12	US-09-795-872-5	Sequence 5, Appl
38	964	16.4	1034	11	US-09-984-130-43	Sequence 43, Appl
39	964	16.4	1034	12	US-09-836-353A-43	Sequence 43, Appl
40	950	16.2	240	15	US-10-102-806-453	Sequence 453, Appl
41	941	16.0	191	11	US-09-805-354-1	Sequence 1, Appl
42	941	16.0	191	15	US-10-144-259-1	Sequence 1, Appl
43	852	14.5	707	9	US-09-764-870-313	Sequence 313, Appl
44	852	14.5	707	15	US-10-125-540-313	Sequence 313, Appl
45	807	13.7	223	12	US-10-002-631C-197	Sequence 197, Appl

ALIGNMENTS

RESULT 1

US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimada, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 100.0%; Score 5879; DB 11; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFCQSVVQLGSRVVVGAPQEIVAANQKGLGYCDYVSTGSCPEI	60
Db	1	FNLDTENAMTFQENARGFCQSVVQLGSRVVVGAPQEIVAANQKGLGYCDYVSTGSCPEI	60
Qy	61	RLOQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQPOPK	120
Db	61	RLOQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQPOPK	120
Qy	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKKSKTLFLSLIMQYSEEF	180

```
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTLFSLMQVSEEF 180
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Db 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Qy 241 TDGEKFGDPLGYEDVPELDRGVIIRVVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 241 TDGEKFGDPLGYEDVPELDRGVIIRVVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Qy 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Qy 361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAIAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAIAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLGAPHYYEOTRGQSVCP 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLGAPHYYEOTRGQSVCP 480
Qy 481 PRQORARWQCDVLYGEGQOPMGRFGAALTVLGDNVNGDKLTDVAICAPGEEDNRGAYVLF 540
Db 481 PRQORARWQCDVLYGEGQOPMGRFGAALTVLGDNVNGDKLTDVAICAPGEEDNRGAYVLF 540
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLSGGQDLTMDGLVDLTGAGHVLRLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLSGGQDLTMDGLVDLTGAGHVLRLRSQ 600
Qy 601 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Db 601 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQPNCLIEDPVSIVLRNLF 720
Db 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQPNCLIEDPVSIVLRNLF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRIFTALFPPEKNCNGNDNICQDDLSITFSFMSLDCLVWG 780
Db 721 SLVGTPLSAFGLNRPVLAEDAQRIFTALFPPEKNCNGNDNICQDDLSITFSFMSLDCLVWG 780
Qy 781 GPREFNVTVRNDEGSDYRTQVTFPPDLDSYBKVSTLQNRORSWRLACESASSTEV 840
Db 781 GPREFNVTVRNDEGSDYRTQVTFPPDLDSYBKVSTLQNRORSWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Db 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Qy 901 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMQHQVSNLQORSUPLISLVFLVPV 960
Db 901 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMQHQVSNLQORSUPLISLVFLVPV 960
Qy 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
Db 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
Qy 1021 FFGIQEEFNATLKGNSLSDWTIKTSHNHLIIVSTAELFNDSDVFTLLPGQAGFVRSQTE 1080
Db 1021 FFGIQEEFNATLKGNSLSDWTIKTSHNHLIIVSTAELFNDSDVFTLLPGQAGFVRSQTE 1080
Qy 1081 KVEPPEVNPPLIVGSSVGLLALLITALYKLGFPKRYKDMSEGPPGAEQ 1137
Db 1081 KVEPPEVNPPLIVGSSVGLLALLITALYKLGFPKRYKDMSEGPPGAEQ 1137
```

RESULT 2:

```
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4
```

```
Query Match 99.6%; Score 5857; DB 11; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 FNLDTENAMTFQBNARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCEPI 60
Db 1 FNLDTENAMTFQBNARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCEPI 60
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQTCSNTYVYKGLCFIFGSLNRQOPQK 120
Db 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQTCSNTYVYKGLCFIFGSLNRQOPQK 120
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTLFSLMQVSEEF 180
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTLFSLMQVSEEF 180
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Db 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Qy 241 TDGEKFGDPLGYEDVPELDRGVIIRVVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 241 TDGEKFGDPLGYEDVPELDRGVIIRVVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Qy 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTQVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Qy 361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAIAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINNTRVDSNMNDAYLGVAIAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLGAPHYYEOTRGQSVCP 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLGAPHYYEOTRGQSVCP 480
Qy 481 PRQORARWQCDVLYGEGQOPMGRFGAALTVLGDNVNGDKLTDVAICAPGEEDNRGAYVLF 540
Db 481 PRQORARWQCDVLYGEGQOPMGRFGAALTVLGDNVNGDKLTDVAICAPGEEDNRGAYVLF 540
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLSGGQDLTMDGLVDLTGAGHVLRLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLSGGQDLTMDGLVDLTGAGHVLRLRSQ 600
Qy 601 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Db 601 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQPNCLIEDPVSIVLRNLF 720
Db 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQPNCLIEDPVSIVLRNLF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRIFTALFPPEKNCNGNDNICQDDLSITFSFMSLDCLVWG 780
```


Db 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Qy 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQONORSQSWRLACESASSTEV 840
Db 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQONORSQSWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIFPENSEVTFTDQVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db 841 SGALKSTSCSINHPIFPENSEVTFTDQVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Qy 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 960
Db 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 960
Qy 961 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1020
Db 961 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1020
Qy 1021 FFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAEILFNDVSFTLLPQGGAFVRSOTET 1080
Db 1021 FFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAEILFNDVSFTLLPQGGAFVRSOTET 1080
Qy 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPPGAEPPQ 1137
Db 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPPGAEPPQ 1137

RESULT 3

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350, 259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193, 043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173, 497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286, 889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362, 652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943, 363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 9; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOQTCSENTYVKGCLFLGSLNRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOQTCSENTYVKGCLFLGSLNRQPOK 136
Qy 121 FPEALRGCCQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCCQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQLKSKTLFSLMQYSEEF 196

RESULT 4

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia

Qy 181 RIHFTFKFQNNPNSRLIKBITQLGRTHATGLRKVYVRELFNITNGARKNAPKILFLL 240
Db 197 RIHFTFKFQNNPNSRLVXPIITQLGRTHATGIRKVVYVRELFNITNGARKNAPKILVVI 256
Qy 241 TDGSKFGDPLGYEDVPELDEGVIRYVVGFDAPRSEKSRQELNVTASKPRDHVFOAN 300
Db 257 TDGSKFGDPLGYEDVPEADREGVIRYVVGFDAPRSEKSRQELNVTASKPRDHVFOAN 316
Qy 301 NFEALKTQONOLREKI PAIEGTOTGSSSSFEHMSQEGFSAATISNGPILLSITVGSYDNAG 360
Db 317 NFEALKTQONOLREKI PAIEGTOTGSSSSFEHMSQEGFSAATISNGPILLSITVGSYDNAG 376
Qy 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 436
Qy 421 QNTGWESNANVKTQICAYFGASLCSVDVDSNGSTDLVLGAPHYYQTRGGQVSVCP 480
Db 437 QNTGWESNANVKTQICAYFGASLCSVDVDSNGSTDLVLGAPHYYQTRGGQVSVCP 496
Qy 481 PRGQARWQCDAVLYGEOQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDAVLYGEOQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSLSGQDLTMDGLVDTLVGAQHVLRLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSLSGQDLTMDGLVDTLVGAQHVLRLRSQ 616
Qy 601 PVLRVKAIEMFNPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTRDLRREGOIQSVVT 660
Db 617 PVLRVKAIEMFNPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTRDLRREGOIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQTCETLKLQPCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQTCETLKLQPCIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADQRLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQONORSQSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLISYRKVSTLQONORSQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTFTDQVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFTDQVDSKASLGNKLLKANVTSENMPRTNKTFF 916
Qy 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 976
Qy 961 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1020
Db 977 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1036
Qy 1021 FFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAEILFNDVSFTLLPQGGAFVRSOTET 1080
Db 1037 FFGIOEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAEILFNDVSFTLLPQGGAFVRSOTET 1096
Qy 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPPGAEPPQ 1137
Db 1097 KVEPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPPGAEPPQ 1153

```

; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.3%; Score 5839; DB 11; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLVQCDSYSGSCEPI 60
Db 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLVQCDSYSGSCEPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGTVHTQCSNTYVKGCLFLGSLNRQOQOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGTVHTQCSNTYVKGCLFLGSLNRQOQOK 136

Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196

Qy 181 RIHTTFKEFQNNPRSLIKPITQLLGRTHATGLRKVKVRELFINITNGARKNAFKILVLI 240
Db 197 RIHTTFKEFQNNPRSLIKPITQLLGRTHATGLRKVKVRELFINITNGARKNAFKILVLI 256

Qy 241 TDGEKFGDPLGYEDVIPPEDREGVIRVYLGFDAPFRSEKSKQELNTVASKPPRHHVQAN 300
Db 257 TDGEKFGDPLGYEDVIPPEDREGVIRVYLGFDAPFRSEKSKQELNTVASKPPRHHVQAN 316

Qy 301 NFEALTKTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSGVSYDWAG 360
Db 317 NFEALTKTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSGVSYDWAG 376

Qy 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436

Qy 421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSVCP 480
Db 437 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSVCP 496

Qy 481 PRGQARWQCDAVLGEOQPMGRFGAALTVLGDNVNGDKLTDAITGAPGEDNRGAVLYF 540
Db 497 PRGQARWQCDAVLGEOQPMGRFGAALTVLGDNVNGDKLTDAITGAPGEDNRGAVLYF 556

Qy 541 HGTSGSGISPSHSQRISAGSLSPRIQYFCQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRISAGSLSPRIQYFCQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616

Qy 601 PVLRYKATMEFNPVARNVECDNVVKGKEAGEVRVCLHVQKSTRDLRREGQIQSVYT 660
Db 617 PVLRYKATMEFNPVARNVECDNVVKGKEAGEVRVCLHVQKSTRDLRREGQIQSVYT 676

Qy 661 VDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 VDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736

Qy 721 SLVGTPLSAFGNLRPVLAEADAQRFLTAFFPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRPVLAEADAQRFLTAFFPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQONORSQSRWLACESASSTEV 840
Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQONORSQSRWLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLLKANVTSENNMPTNKTEF 916

Qy 901 QLELPVKYAYVMVVTSHGVSTKYLNTASENTSRVMQHOYQVSNLQORSPLISLVLPV 960
Db 917 QLELPVKYAYVMVVTSHGVSTKYLNTASENTSRVMQHOYQVSNLQORSPLISLVLPV 976

Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1036

Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDVSFTLLPGQGAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDVSFTLLPGQGAFVRSQTET 1096

Qy 1081 KVPEFVNPDLPLIVGSSVGGILLLALITAAALKLGFKKQYKDMHSEGGPPGAEPQ 1137
Db 1097 KVPEFVNPDLPLIVGSSVGGILLLALITAAALKLGFKKQYKDMHSEGGPPGAEPQ 1153

RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.3%; Score 5839; DB 11; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLVQCDSYSGSCEPI 60
Db 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVAAANQSGSLVQCDSYSGSCEPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGTVHTQCSNTYVKGCLFLGSLNRQOQOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGTVHTQCSNTYVKGCLFLGSLNRQOQOK 136

Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPDFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196

Qy 181 RIHTTFKEFQNNPRSLIKPITQLLGRTHATGLRKVKVRELFINITNGARKNAFKILVLI 240
Db 197 RIHTTFKEFQNNPRSLIKPITQLLGRTHATGLRKVKVRELFINITNGARKNAFKILVLI 256
```

```
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAFRSEKSRQBLNTVASKPRDHFVQAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVLGFGDAFRSEKSRQBLNTVASKPRDHFVQAN 316
QY 301 NFALATVQNLREKIFAIEGTGTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDMAG 360
Db 317 NFALATVQNLREKIFAIEGTGTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDMAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLVLAGPQYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLVLAGPQYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 496
QY 481 PRQQRARWOCDAVLXGEGQPMGRFGAALTVDLVGNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQQRARWOCDAVLXGEGQPMGRFGAALTVDLVGNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHRSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHRSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNSRRQTQVLGTLTQTCETIKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSRRQTQVLGTLTQTCETIKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRFTALPFPEKNCNGNDNICQDDLSTITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALPFPEKNCNGNDNICQDDLSTITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQONORSQRWSRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQONORSQRWSRLACESASSTEV 856
```

RESULT 6

```
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
```

```
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-259-30
```

```
Query Match 99.3%; Score 5839; DB 15; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAFQEIIVAANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAFQEIIVAANORGSLYQCDYSTGSCPEI 76
QY 61 RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFGLFGSNLRQPOK 120
Db 77 RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFGLFGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMOYSEEF 196
QY 181 RIHFTKFEQNNPNRSLIKPITQLLGRTHATATGLRKVVRELFNITNGARKNAFKILFL 240
Db 197 RIHFTKFEQNNPNRSLIKPITQLLGRTHATATGLRKVVRELFNITNGARKNAFKILV 256
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAFRSEKSRQBLNTVASKPRDHFVQAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVLGFGDAFRSEKSRQBLNTVASKPRDHFVQAN 316
QY 301 NFEALKTQVQNLREKIFAIEGTGTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDMAG 360
Db 317 NFEALKTQVQNLREKIFAIEGTGTGSSSFEHEMSQEGFSAAITSNGLPSTVGSYDMAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLVLAGPQYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLVLAGPQYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 496
QY 481 PRQQRARWOCDAVLXGEGQPMGRFGAALTVDLVGNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQQRARWOCDAVLXGEGQPMGRFGAALTVDLVGNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHRSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHRSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNSRRQTQVLGTLTQTCETIKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSRRQTQVLGTLTQTCETIKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRFTALPFPEKNCNGNDNICQDDLSTITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALPFPEKNCNGNDNICQDDLSTITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQONORSQRWSRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQONORSQRWSRLACESASSTEV 856
```

QY 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKLLKANVTSENMMPTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKLLKANVTSENMMPTNKTEF 916
QY 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHQYQVSNLQORSPLPSLFLVPV 960
DB 917 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHQYQVSNLQORSPLPSLFLVPV 976
QY 961 RLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1020
DB 977 RLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1036
QY 1021 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGOGAFVRSOTET 1080
DB 1037 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGOGAFVRSOTET 1096
QY 1081 KVEPFEVNPPLPIVGSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
DB 1097 KVEPFEVNPPLPIVGSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 7
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207.655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.3%; Score 5839; DB 15; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVWOLQGSRRVVGAPQEIIVAAQORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARFGQSVWOLQGSRRVVGAPQEIIVAAQORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVKGCLCFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVKGCLCFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQSDSIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLPFLMOYSEEP 180
DB 137 FPEALRGCPQSDSIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLPFLMOYSEEP 196
QY 181 RIHTFKEFQNNPRSLIKBITQLGRTHATGLRKVVRELFNITGARKNAKILFL 240
DB 197 RIHTFKEFQNNPRSLIKBITQLGRTHATGLRKVVRELFNITGARKNAKILVVI 256
QY 241 TDGKEFGDPLGVEDYVPELDRGVIRYVVGDAFRSEKSEQLNTVASKPPRDHVFOAN 300
DB 257 TDGKEFGDPLGVEDYVPEADREGVIRYVVGDAFRSEKSEQLNTVASKPPRDHVFOAN 316
QY 301 NFEALKTQNLREKIPAEIGTQTGSSSSFFHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
DB 317 NFEALKTQNLREKIPAEIGTQTGSSSSFFHEMSQEGFSAAITSNGLPLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTVDSDMDAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVDSDMDAYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480

DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDVAVLYGEOGPWRFGAALTVLGDVYNGDKLTDAVAGAPGEENRGAVYLP 540
DB 497 PRGQARWQCDVAVLYGEOGPWRFGAALTVLGDVYNGDKLTDAVAGAPGEENRGAVYLP 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFCQSLSGGGQDLTMDGLVDLTVGAQGHVLLLRQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFCQSLSGGGQDLTMDGLVDLTVGAQGHVLLLRQ 616
QY 601 PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLRGRQIQSVVT 660
DB 617 PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLRGRQIQSVVT 676
QY 661 YDLALDSGPHSRVAFNETKSTRQTOVLGTTTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGPHSRVAFNETKSTRQTOVLGTTTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAORLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNLRPVLAEDAORLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACSSASTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACSSASTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKLLKANVTSENMMPTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKLLKANVTSENMMPTNKTEF 916
QY 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHQYQVSNLQORSPLPSLFLVPV 960
DB 917 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHQYQVSNLQORSPLPSLFLVPV 976
QY 961 RLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1020
DB 977 RLNQTVMWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1036
QY 1021 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGOGAFVRSOTET 1080
DB 1037 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGOGAFVRSOTET 1096
QY 1081 KVEPFEVNPPLPIVGSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
DB 1097 KVEPFEVNPPLPIVGSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 8
US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902.481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-5

Query Match 99.3%; Score 5836; DB 11; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOQPK 120
DB 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOQPK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRFMKELVSTIMEQLKSKTFLSLMQSYSEF 180
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRFMKELVSTIMEQLKSKTFLSLMQSYSEF 180

QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240
DB 181 RHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQAN 300
DB 241 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQAN 300

QY 301 NFEALKTQONLREKIPAIETGQTGSSSFEHEMSQEGFSAITNSGPLLSTVSGSDWAG 360
DB 301 NFEALKTQONLREKIPAIETGQTGSSSFEHEMSQEGFSAITNSGPLLSTVSGSDWAG 360

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYQTRGGQSVSCPL 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYQTRGGQSVSCPL 480

QY 481 PRGQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

QY 541 HGTSGSGISPSHSORTAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVCAGQHVLLRSQ 600
DB 541 HGTSGSGISPSHSORTAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVCAGQHVLLRSQ 600

QY 601 PVLVRKAIINEFPREVARNVFCNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLVRKAIINEFPREVARNVFCNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660

QY 661 YDLALDSGRPHSRAVENETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 661 YDLALDSGRPHSRAVENETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720

QY 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFFKNCNDNI CODDLSITFSFMSLCLVVG 780
DB 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFFKNCNDNI CODDLSITFSFMSLCLVVG 780

QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPLDLSYRKVSTLQONRSQSRWLACESASSTEV 840
DB 781 GPREFNVTVVRNDGEDSVRTQVTFPPLDLSYRKVSTLQONRSQSRWLACESASSTEV 840

QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900

QY 901 QLELPVKYAVMVVTSHGVSSTKYLNFPTASENTRVMQHOYQVSNLQORSPLPSLFLVPV 960
DB 901 QLELPVKYAVMVVTSHGVSSTKYLNFPTASENTRVMQHOYQVSNLQORSPLPSLFLVPV 960

QY 961 RLNQTVIWRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCISIAVCQRIOCDIP 1020
DB 961 RLNQTVIWRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCISIAVCQRIOCDIP 1020

QY 1021 PFGIOEEFNATLKGNSLDFDWIKTSHNHLIIVSTAAILFNDSVFTLLPQGGAFVRSOTET 1080

DB 1021 PFGIOEEFNATLKGNSLDFDWIKTSHNHLIIVSTAAILFNDSVFTLLPQGGAFVRSOTET 1080

QY 1081 KVEPEVEVNPPLIIVGSSVGGLLALITALYKLGFFKQYKDMWSEGPPGASBPQ 1137
DB 1081 KVEPEVEVNPPLIIVGSSVGGLLALITALYKLGFFKQYKDMWSEGPPGASBPQ 1137

RESULT 9
US-09-902-481A-6
; Sequence 6; Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RWK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.2%; Score 5832; DB 11; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOQPK 120
DB 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOQPK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRFMKELVSTIMEQLKSKTFLSLMQSYSEF 180
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRFMKELVSTIMEQLKSKTFLSLMQSYSEF 180

QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240
DB 181 RHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQAN 300
DB 241 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQAN 300

QY 301 NFEALKTQONLREKIPAIETGQTGSSSFEHEMSQEGFSAITNSGPLLSTVSGSDWAG 360
DB 301 NFEALKTQONLREKIPAIETGQTGSSSFEHEMSQEGFSAITNSGPLLSTVSGSDWAG 360

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYQTRGGQSVSCPL 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYQTRGGQSVSCPL 480

QY 481 PRGQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGOSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSORIASGLSPRLQYFGOSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 600
QY 601 PVLKVAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIOSVVT 660
DB 601 PVLKVAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIOSVVT 660
QY 661 YDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQPNCLIEDPVSPIVLRNF 720
DB 661 YDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQPNCLIEDPVSPIVLRNF 720
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRYKSVSTLQNRQSQRWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRYKSVSTLQNRQSQRWRLACESASSTEV 840
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
QY 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOVQVSNLQORSIPISLVFLVPV 960
DB 901 QLELPVKYAVYVMVTSYKYNFTASENTSRVMQHOVQVSNLQORSIPISLVFLVPV 960
QY 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
DB 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
QY 1021 FPGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
DB 1021 FPGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
QY 1081 KVEPFEVNPPLPLIVGSSVGLLLALITAAALKYKLGFFKQYKDMSEGGPPGAEPQ 1137
DB 1081 KVEPFEVNPPLPLIVGSSVGLLLALITAAALKYKLGFFKQYKDMSEGGPPGAEPQ 1137
RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945.265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4
Query Match 99.1%; Score 5823.5; DB 10; Length 1152;
Best Local Similarity 98.8%;
Matches 1123; Conservative 8; Mismatches 5; Indels 1; Gaps 1;
QY 1 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAQNRCSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAQNRCSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVVKGLCFGLGNSLRQPOK 120

RESULT 11


```
Db 138 LPVSRQCPROEQDIIFLIDSGSISRNFAFMVRAVISQFQRPSTQFSLMQSNKF 197
Qy 181 RIHPTFEPQNNPRSLIKPIIQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 198 QTHPTFEPFRRTSNPLSLASVHQLQGFTYATAIQNVVHRLFHASYGARRDAIKILIVI 257
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSOELNTVASKPPRDHVFQAN 300
Db 258 TDGKKGSDLDYKDVIPMADAAGIIRYAGVGLAFQNRNSWKEINDIASPSQEHFKVE 317
Qy 301 NFEALKTQVQUREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 318 DFDALKDIONQLEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGVLGAVGSFTWSG 377
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVSQSLVGLGAPRYQHIGKAVIFI 437
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCPPL 480
Db 438 QVSRQWRMKAIEVIGTIGSYFGASLCSVDVDTGSDTDLVIGAPHYYEOTRGGOVSVCPPL 497
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 498 PRGWR-RWMCDAVLYGEGQHPWGRFGAALTVDVNGDKLTDVVI GAPGEENRGAYVLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQALSGQDLTQDGLVDLAVGARGVLLLRTR 616
Qy 601 PVLVRKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLVHVKSTDRREGIOISVVT 660
Db 617 PVLWVGVMQFIPAEIPRSAFECEQVVSQTLVQSNICILYIDKRSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQPNCTEDVPSTVLRNLF 720
Db 677 LDLALAPGRLSPRAIFOETKNRSLSRVRLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736
Qy 721 SLVQTPISAFQNLRPVLAEDAQRULTALFPFEKNCQNDNTCCQDLSTFTFSFMSLDCULVG 780
Db 737 TLVGKPLLAFLNLRPLMALLAQAQRYFTASLPFEKNCQADHICQDNLGIFSFPGLKSLVG 796
Qy 781 GPREFNVTVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRMSRLACESASTEV 840
Db 797 SNLEINAEVMVMDGEDSYGTTITFSPAGLSYRYVAEGQKQGLRSLHLC--CSAPVG 854
Qy 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLILKANVTSENNMPRTNKTEF 900
Db 855 SQGTWSTSCRINHILFRGGAQITFLATFDVSPKAVGLDRULLLIANVSSENNIPTSKTIF 914
Qy 901 QLELPVKYAVVMVTSYHGVSTKYLNFTAS-ENTSRVMQHOVQVSNLQORSPLISLFLVP 959
Db 915 QLELPVKYAVYIVVSSHQETKYLNFSSEKESHVAMHRYQVNNLQORDLPVSNFNPWP 974
Qy 960 VRLNQTVMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORICDI 1019
Db 975 VELNQEAVMMDVEVSHQNPISRCSEKSIAPPADFLAHQKNPVLDCSTAGCLRFRCDV 1034
Qy 1020 PFFGQBEFNATLKNLSFDWYIKTSHNHLIVSTAELFNDSVFTLLPGOGAFVRQTE 1079
Db 1035 PPSVQBELEDFTLKNLSFGHVRQILOKKSVSVAEIIIFDTSVYSQLPQGEAFMRAQTI 1094
Qy 1080 TKVPEFVFNPLPIVSGSVGGLLLALITAAALYKLGFFKRYQKDMXSE 1128
Db 1095 TVLEKIKVNDIPLIVSGSIGGLLLALITAVLYKVGFFKRYQKEMMEE 1143
```

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US20030077278A1

; GENERAL INFORMATION:

```
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-4
```

Query Match 58.2%; Score 3423; DB 11; Length 1163;

Best Local Similarity 60.2%; Pred. No. 3e-307;

Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

```
Qy 1 FNLDTENAMTPOENARGFCQSVVQLQGSRRVVVGAPQEIIVAAQNRGSLYOCDSYSGSCEPI 60
Db 20 FNLDTBELTAFRVDVAGFDGVVQVANSVVVGAPQKIITAAQIGLYOCYSTGACEPI 79
Qy 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVYKGLCFGLGSLNRQPOK 120
Db 80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRNMVLTGLCFLLGPT--QLTOR 137
Qy 121 FPEALRGCEQSDIAFLIDGSGSIIPHPFRMKELVSTIMEQLKSKTLFSLMQVSEEF 180
Db 138 LPVSRQCPROEQDIIFLIDGSGSISRNFAFMVRAVISQFQRPSTQFSLMQSNKF 197
Qy 181 RTHFTKEFQNNPNPSLIKPIIQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 198 QTHPTFEPFRRTSNPLSLASVHQLQGFTYATAIQNVVHRLFHASYGARRDAIKILIVI 257
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSOELNTVASKPPRDHVFQAN 300
Db 258 TDGKKGSDLDYKDVIPMADAAGIIRYAGVGLAFQNRNSWKEINDIASPSQEHFKVE 317
Qy 301 NFEALKTQVQUREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 318 DFDALKDIONQLEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGVLGAVGSFTWSG 377
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVSQSLVGLGAPRYQHIGKAVIFI 437
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCPPL 480
Db 438 QVSRQWRMKAIEVIGTIGSYFGASLCSVDVDTGSDTDLVIGAPHYYEOTRGGOVSVCPPL 497
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 498 PRGWR-RWMCDAVLYGEGQHPWGRFGAALTVDVNGDKLTDVVI GAPGEENRGAYVLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQALSGQDLTQDGLVDLAVGARGVLLLRTR 616
Qy 601 PVLVRKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLVHVKSTDRREGIOISVVT 660
Db 617 PVLWVGVMQFIPAEIPRSAFECEQVVSQTLVQSNICILYIDKRSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQPNCTEDVPSTVLRNLF 720
Db 677 LDLALAPGRLSPRAIFOETKNRSLSRVRLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736
```


APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. US2003007278A1el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/891,943
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1161
TYPE: PRN
ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 57.6%; Score 3388; DB 11; Length 1161;
Best Local Similarity 58.9%; Pred. No. 5.3e-304;
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

```
Qy 1 FNLDTENAMTQENARGEGOSVQLOGRVVGAPQEIIVAAANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDOVEETIFQEDAGGFGQSVQFGGSRVVGAPLEVVAANQGRLYDCAATGMCQPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATSPOLLACGTPVHTQCSNTYVKGCLGFLGSGNLRQOPQK 120
Db 77 PLHTRPEAVNMSLGLTAASTNGSRLLACGTPLHRVCGENSYSKGSCLLGSRW-ELIQT 135

Qy 121 FPEALRGCPQSDSIAPLIDGSGIIPHDPRMKELVSTIMEQKSKTFLSLMOYSEEF 180
Db 136 VPDATPECPHOEMDIVFLIDGSGIDQNDQNMKGFGVQAVNMGQFEGDTFLPALMOYSNLL 195

Qy 181 RIHTEFEFQNNPNSLIKPITOLLGRTHATGLRKVVRLEFNITNGARKNAFKILFL 240
Db 196 KIHFTFTQRTSPQOSSLVDPLVQLKGLTFTATGLTIVTQFHHKNGARKSAKKILIVI 255

Qy 241 TDGEKFGDPLGYEDVPELDEGRIYVVLGFGDAPRSEKQELNTVASKPPRDHVFOAN 300
Db 256 TDGQKYKDPLEYSVDVIPAERAGIIRYAIVGHAFOGPTARQELNTISSAPPQDHFVKVD 315

Qy 301 NFEALKTVQNLREKIFAIETGTQSSSFPEHMSQEGFSAITSNGPLLTGVSQYDWAG 360
Db 316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGSFWSG 375

Qy 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 376 GAFLYPPNNSPTFINMSQENVDMRDSYLGYSTELAMKGVQNLVLGAPRYOHTKAVIFT 435

Qy 421 QNTGWESNANVKGITGAYFGASLCSVDVDSNGSTDVLVLCAPHYEQTGGQVSVCP 480
Db 436 QVSRQWRKKAETVGTIGSYFGASLCSVDVDSGSTDLLIGAPHYEQTGGQVSVCP 495

Qy 481 PRGQARWQCDVAVYGEQCPMGRFGAALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 540
Db 496 PRGQVQWQCDVAVRGEQHPMGRFGAALTVLGDVNEKLIIDVAIGAPCEQENRGAVYLF 555

Qy 541 HCTSGSGLSPSHSRIAGSKLSPRLQYFGOSLGGODLTMDGLVDLTGAGQGHVLLRSQ 600
Db 556 HGASEGSLSPSHSRIAGSKLSPRLQYFGQALSGGODLTQDGLMDLAVGARGQVLLRLSL 615

Qy 601 PVLRYKAIEMFNPVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 616 PVLKGVAMRFSFVEVAKAVYKWEKPSALEAGATVCLTIQKSSLDQL--GDIQSVR 673

Qy 661 YDLALDSGRPHSRVFNFTKNSRRQTQVLGTQTCETLKLQLPNCIEDPVPPIVRLNF 720
Db 674 FDLALDPGRLTSAIFNETKNTPLTRKTLGLGHCETLKLALLPCDVEDVVSPIILHLNF 733
```

```
Qy 721 SLVGTPLSAPGNLRPVLAEQAQLFTALFPFPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db 734 SLVREPIPSQNLRLPVLAVGSQDLFTASLPFPEKNCGQDGLCEGDLGVTLSFSGLTQTLTVG 793

Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSVRKYUSTLQNBQSRQSWRLACESASSTEV 840
Db 794 SSLELNVIIVTMNAGEDSYGTVVSLYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852

Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANKVTSENNMPTNKTEF 900
Db 853 EG-LRSRCSVNHPIFHEGSGNGTFIVTFDVSYKATLGDRMLMRASASSENKASSKATF 911

Qy 901 QLELPVKYAVYVMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPLISLVLVP 959
Db 912 QLELPVKYAVYTMISRQESTKYFNATSDKKMKAEHRYRVNLSQRDLAISINFWVP 971

Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDI 1019
Db 972 VLLNGVAVMDVVMNEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQPCDV 1029

Qy 1020 PFFGIOBEFNATLKGNLSPDWYIKTSHNHLIIVSTAELFNDSVFTLLPQCGAFVRSQTE 1079
Db 1030 PFSVQVEELDFTLKGNLSFGWVRETLOKKVLVSVVAEITFDTSVYQOLPQOEAFMAQME 1089

Qy 1080 TKYEPPEVPNPLPLIVGSSVGGLLLLALITALIYALYKLGFFKRYKQYKMMSE 1128
Db 1090 MVLEDEEVYNAIPIIMGSSVYVALLLALITATLYKLGFFKRYKHEMLED 1138
```

Search completed: November 25, 2003, 14:41:11
Job time : 31.1265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using ew model

Run on: November 25, 2003, 14:13:09 ; Search time 13.5618 Seconds
(without alignments)
3547.268 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPGGABQP 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*

2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*

3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*

4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*

5: /cgn2_6/prodata/1/iaa/6CTUS_COMB.pep:*

6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	1	US-08-173-497-3
2	5839	99.3	1153	1	Sequence 3, Appli
3	5839	99.3	1153	1	Sequence 3, Appli
4	5839	99.3	1153	1	Sequence 3, Appli
5	5839	99.3	1153	2	Sequence 3, Appli
6	5839	99.3	1153	2	Sequence 3, Appli
7	5839	99.3	1153	2	Sequence 3, Appli
8	5839	99.3	1153	3	Sequence 3, Appli
9	5839	99.3	1153	4	Sequence 3, Appli
10	5808.5	98.8	1152	5	Sequence 43, Appli
11	5808.5	98.8	1152	5	Sequence 43, Appli
12	5808.5	98.8	1152	6	Patent No. 5424399
13	3446	58.6	1163	5	US-08-476-062A-44
14	3446	58.6	1163	5	PCT-US96-01314-44
15	3423	58.2	1163	1	US-08-173-497-4
16	3423	58.2	1163	1	Sequence 4, Appli
17	3423	58.2	1163	1	Sequence 4, Appli
18	3423	58.2	1163	1	Sequence 4, Appli
19	3423	58.2	1163	2	Sequence 4, Appli
20	3423	58.2	1163	2	US-08-605-672-4
21	3423	58.2	1163	2	US-08-482-293A-4
22	3423	58.2	1163	3	US-08-943-363-4
23	3423	58.2	1163	4	US-09-193-043-4
24	3388	57.6	1161	1	US-08-173-497-2
25	3388	57.6	1161	1	Sequence 2, Appli
26	3388	57.6	1161	1	Sequence 2, Appli
27	3388	57.6	1161	1	Sequence 2, Appli

28 3388 57.6 1161 2 US-08-605-672-2 Sequence 2, Appli
29 3388 57.6 1161 2 US-08-482-293A-2 Sequence 2, Appli
30 3388 57.6 1161 3 US-08-943-363-2 Sequence 2, Appli
31 3388 57.6 1161 3 US-09-193-043-2 Sequence 2, Appli
32 3388 57.6 1161 4 US-09-688-307A-2 Sequence 2, Appli
33 3372.5 57.4 1161 1 US-08-485-618-99 Sequence 99, Appli
34 3372.5 57.4 1161 2 US-08-605-672-99 Sequence 99, Appli
35 3372.5 57.4 1161 2 US-08-482-293A-99 Sequence 99, Appli
36 3372.5 57.4 1161 2 US-08-943-363-99 Sequence 99, Appli
37 3372.5 57.4 1161 3 US-09-193-043-99 Sequence 99, Appli
38 3372.5 57.4 1161 3 US-09-688-307A-99 Sequence 99, Appli
39 3213.5 54.7 1161 3 US-09-193-043-55 Sequence 55, Appli
40 3213.5 54.7 1161 4 US-09-688-307A-55 Sequence 55, Appli
41 3206.5 54.5 1161 1 US-08-485-618-55 Sequence 55, Appli
42 3206.5 54.5 1161 1 US-08-362-652-55 Sequence 55, Appli
43 3206.5 54.5 1161 2 US-08-605-672-55 Sequence 55, Appli
44 3206.5 54.5 1161 2 US-08-482-293A-55 Sequence 55, Appli
45 3206.5 54.5 1161 2 US-08-943-363-55 Sequence 55, Appli

ALIGNMENTS

RESULT 1

US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-173-497-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEI VAAHQGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEI VAAHQGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLGSLNRQOQPK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLGSLNRQOQPK 136
QY 121 FPEALRCQPEDSDIAFLIDGSGSIIPHDFRMEKELVSTIMEQLKSKTFLFSLMQYSEEP 180
DB 137 FPEALRCQPEDSDIAFLIDGSGSIIPHDFRMEKELVSTIMEQLKSKTFLFSLMQYSEEP 196
QY 181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVRKVFNITNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVRKVFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDRREGVIRYVLGDFGAFRSEKSRQELNTVASKPRDRHVFQAN 300
DB 257 TDGEKFGDPLGYEDVPELDRREGVIRYVLGDFGAFRSEKSRQELNTVASKPRDRHVFQAN 316
QY 301 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 360
DB 317 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTVDSDMDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVDSDMDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDLVLGAPHYETQTRGGVSVCP 480
DB 437 QNTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDLVLGAPHYETQTRGGVSVCP 496
QY 481 PRGARARQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
DB 497 PRGARARQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKATMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
DB 617 PVLRVKATMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 676
QY 661 YDLALDGRPHSRVFNTEKSTRQTVLGLTQTCETLKLQPNCTIEDPVPISVLRLNF 720
DB 677 YDLALDGRPHSRVFNTEKSTRQTVLGLTQTCETLKLQPNCTIEDPVPISVLRLNF 736
QY 721 SLVGTPLSAFGLNRLPVLAEADQRLFTALFPPEKNCNDNIQDDLSITFSPMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNRLPVLAEADQRLFTALFPPEKNCNDNIQDDLSITFSPMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 916
QY 901 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHOYQVSNLQBSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHOYQVSNLQBSLPISLVFLVPV 976
QY 961 RLNQTVIWDROPVTFSENLSSCTCHTKERLPHSDFLAELRKAPVNCISIAVCQICDIP 1020
DB 977 RLNQTVIWDROPVTFSENLSSCTCHTKERLPHSDFLAELRKAPVNCISIAVCQICDIP 1036
QY 1021 PFGIQEEFNATLKGNSLDWIKTSHNHLILVSTAEILFNDPSVFTLLPGQAFVRSQTE 1080
DB 1037 PFGIQEEFNATLKGNSLDWIKTSHNHLILVSTAEILFNDPSVFTLLPGQAFVRSQTE 1096
QY 1081 KVEPEFVNPLPIIVGSSVGGLLLLALITAALYKLGFFKQYKDMMSSEGGPPGABPQ 1137
DB 1097 KVEPEFVNPLPIIVGSSVGGLLLLALITAALYKLGFFKQYKDMMSSEGGPPGABPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGSLNRQOQPK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGSLNRQOQPK 136
QY 121 FPEALRCQPEDSDIAFLIDGSGSIIPHDFRMEKELVSTIMEQLKSKTFLFSLMQYSEEP 180
DB 137 FPEALRCQPEDSDIAFLIDGSGSIIPHDFRMEKELVSTIMEQLKSKTFLFSLMQYSEEP 196
QY 181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVRKVFNITNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVRKVFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDRREGVIRYVLGDFGAFRSEKSRQELNTVASKPRDRHVFQAN 300
DB 257 TDGEKFGDPLGYEDVPELDRREGVIRYVLGDFGAFRSEKSRQELNTVASKPRDRHVFQAN 316
QY 301 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 360
DB 317 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 376

```
QY 361 GVFLYTSKEKSTFINNTRVDSNDNDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSNDNDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQGVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQGVSVCP 496
QY 481 PRGQARWQCCDAVLYGEOGPNGRGAALTVDVGVNGDKLTDVAIGAGEEDNRCNAVYLF 540
DB 497 PRGQARWQCCDAVLYGEOGPNGRGAALTVDVGVNGDKLTDVAIGAGEEDNRCNAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSLGSGGQDLTMDGLVLTGVAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSLGSGGQDLTMDGLVLTGVAQGHVLLRSQ 616
QY 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGPHSRVFNETHNSTRTQVLTGTLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGPHSRVFNETHNSTRTQVLTGTLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDEGDSVRTQVTFPPPLDLSYRVKSTLQNRQSRWRLACASSTEV 840
DB 797 GPREFNVTVVRNDEGDSVRTQVTFPPPLDLSYRVKSTLQNRQSRWRLACASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVFNITPDVDSKASLGNKLLIKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVFNITPDVDSKASLGNKLLIKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVWVTSHGVSITKYNFTASENTSRVWQHGYQVSNLQSRSLPISLFLVPV 960
DB 917 QLELPVKYAVVWVTSHGVSITKYNFTASENTSRVWQHGYQVSNLQSRSLPISLFLVPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPVWNCSTAVCORIQCDIP 1020
DB 977 RLNQTVIMDRPQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPVWNCSTAVCORIQCDIP 1036
QY 1021 PFGIOBEFNATLKNLSFQWYIKTSHNHLIIVSTABILFNDVSFTLLPQGGAFVRSQTEF 1080
DB 1037 PFGIOBEFNATLKNLSFQWYIKTSHNHLIIVSTABILFNDVSFTLLPQGGAFVRSQTEF 1096
QY 1081 KVEPPEVPNPLPLIVGSSVGGLLLALITAAALYKLGFFKRYKDMWSEGGPPGAEPQ 1137
DB 1097 KVEPPEVPNPLPLIVGSSVGGLLLALITAAALYKLGFFKRYKDMWSEGGPPGAEPQ 1153
```

RESULT 3

```
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173.497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286.889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362.652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-3
```

```
Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPEIIVAAANQRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPEIIVAAANQRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNNSLGLSLAATTSPQLLACGTTHQTCSENTYVKGCLCFGLSGNLRLQOQPK 120
DB 77 RLQVPVEAVNNSLGLSLAATTSPQLLACGTTHQTCSENTYVKGCLCFGLSGNLRLQOQPK 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSII PHDFRRMKELVSTIMEOLKSKTFLPSLMQYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLIDGSGSII PHDFRRMKELVSTIMEOLKSKTFLPSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATATGLRKVVRELFTNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATATGIRKVVRELFTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAPFRSEKSRQELMTVASKPPRDHVFQAN 300
DB 257 TDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAPFRSEKSRQELMTVASKPPRDHVFQAN 316
QY 301 NFEALKTVQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVSGSYDWAG 360
DB 317 NFEALKTVQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVSGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSNDNDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSNDNDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQGVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQGVSVCP 496
QY 481 PRGQARWQCCDAVLYGEOGPNGRGAALTVDVGVNGDKLTDVAIGAGEEDNRCNAVYLF 540
DB 497 PRGQARWQCCDAVLYGEOGPNGRGAALTVDVGVNGDKLTDVAIGAGEEDNRCNAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSLGSGGQDLTMDGLVLTGVAQGHVLLRSQ 600
```

Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGCQDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLRVKAIMBNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVWT 660
Db 617 PVLRVKAIMBNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVWT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFNGLRPVLAEADAQRLLFTALFFPEKNCNDNICQDDLSITFFSMJLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEADAQRLLFTALFFPEKNCNDNICQDDLSITFFSMJLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRACASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRACASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTBF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTBF 916
Qy 901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPV 976
Qy 961 RLNQTVDWRDQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1020
Db 977 RLNQTVDWRDQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1036
Qy 1021 PFGIQEBFNATLKNLGFWDWIKTSHNHLIVSTAELFNDSDVFTLLPGOGAFVRSQTET 1080
Db 1037 PFGIQEBFNATLKNLGFWDWIKTSHNHLIVSTAELFNDSDVFTLLPGOGAFVRSQTET 1096
Qy 1081 KVEPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1153

RESULT 4

US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3
Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGCEPI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOK 136
Qy 121 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKTLFSLMOYSEEF 180
Db 137 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKTLFSLMOYSEEF 196
Qy 181 RIHFTKEFQNNPNRSLKPIITQLLGRTHRTATGLRKVVRLEFNITNGARKNAFKILFL 240
Db 197 RIHFTKEFQNNPNRSLKPIITQLLGRTHRTATGLRKVVRLEFNITNGARKNAFKILV 256
Qy 241 TDGEKFGDPLGYEDVTPELDREGVIRYVLGFGDAPRSEKSRQBLNTVASKPPRDHVPQAN 300
Db 257 TDGEKFGDPLGYEDVTPELDREGVIRYVLGFGDAPRSEKSRQBLNTVASKPPRDHVPQAN 316
Qy 301 NPEALKTVQNLREKJFAIEGTOTGSSSFEHEMSOEGFSAAITSNGLPSTVGSVDWAG 360
Db 317 NFEALKTIQNLREKJFAIEGTOTGSSSFEHEMSOEGFSAAITSNGLPSTVGSVDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGVAIAIILNRNVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGVAIAIILNRNVQSLVLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCP 496
Qy 481 PRGORARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRRGAYL 540
Db 497 PRGORARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRRGAYL 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGCQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGCQDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLRVKAIMBNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVWT 660
Db 617 PVLRVKAIMBNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVWT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFNGLRPVLAEADAQRLLFTALFFPEKNCNDNICQDDLSITFFSMJLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEADAQRLLFTALFFPEKNCNDNICQDDLSITFFSMJLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRACASSTEV 840

Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSVRKUSTLQNSQRSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1036
Qy 1021 FFGIQEENATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSFTLLPQGAFVRSQTET 1080
Db 1037 FFGIQEENATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSFTLLPQGAFVRSQTET 1096
Qy 1081 KVEPFVNPPLPLIVGSSVGLLLALITAAALYKLGFFKRYKDMSEGGPGABEQ 1137
Db 1097 KVEPFVNPPLPLIVGSSVGLLLALITAAALYKLGFFKRYKDMSEGGPGABEQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-605-672-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVVOLOGSRVVVGAPQEIVAANQORSLQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVVOLOGSRVVVGAPQEIVAANQORSLQCDYSTGSCBPI 76
Qy 61 RIQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLFSGNLNQPOPK 120
Db 77 RIQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLFSGNLNQPOPK 136
Qy 121 FPEARLGCQEDSDIAFLIDSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMQYSEB 180
Db 137 FPEARLGCQEDSDIAFLIDSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMQYSEB 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVVGDAFRSEKSRQELNTVASKPPDRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVPELDREGVIRYVVGDAFRSEKSRQELNTVASKPPDRDHVFOAN 316
Qy 301 NFEALKTQVQNLREKIPAEIGTQTCSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 317 NFEALKTQVQNLREKIPAEIGTQTCSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAIIILNRVQSLVGLGAPYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAIIILNRVQSLVGLGAPYQHIGLVAMPR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYETRGQVSVCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYETRGQVSVCP 496
Qy 481 PRGQARMQCDVAVLYGEOQPMGRFGAALTVDLVNGDKLTDAIAGAPGEEDNRGAVYLF 540
Db 497 PRGQARMQCDVAVLYGEOQPMGRFGAALTVDLVNGDKLTDAIAGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTGVAQGHVLLLRQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTGVAQGHVLLLRQ 616
Qy 601 PVLRYKATMEFPNPREVARNVFCNQVVKGEAGEVVRCLHVQKSTRDRRLRSGQSVVT 660
Db 617 PVLRYKATMEFPNPREVARNVFCNQVVKGEAGEVVRCLHVQKSTRDRRLRSGQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALPPFPKNGCNDNICODDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADQRLFTALPPFPKNGCNDNICODDLSITFSFMSLCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSVRKUSTLQNSQRSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSVRKUSTLQNSQRSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMHQYQVSNLQORSLSPLSLVFLVPV 976
Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1036

QY 1021 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
DB 1037 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1096
QY 1081 KVEPFEVNPPLPIVSGVGLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1137
DB 1097 KVEPFEVNPPLPIVSGVGLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1153

RESULT 6

US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams JK., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPEIQAANQRGSLYQCDYSGSCEPI 60
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPEIQAANQRGSLYQCDYSGSCEPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFPLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFPLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMBOLKSKTFLSLMQYSEEF 180

DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMBOLKSKTFLSLMQYSEEF 196
QY 181 RHFTPFKEQNNPNRSLIKPIITOLLGRTHATGLRKVVRELFNITNGARKNAFKTLFLL 240
DB 197 RHFTPFKEQNNPNRSLIKPIITOLLGRTHATGLRKVVRELFNITNGARKNAFKTLVVI 256
QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNLTVAASKPRDHRVFOAN 300
DB 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNLTVAASKPRDHRVFOAN 316
QY 301 NFEALKTQVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLLSTVGSVDWAG 360
DB 317 NFEALKTQVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLLSTVGSVDWAG 376
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSILVLCAPYOHIGLVAMER 420
DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSILVLCAPYOHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCPL 496
QY 481 PRGORARWQCDVLYGEGQOPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYLFP 540
DB 497 PRGORARWQCDVLYGEGQOPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAYLFP 556
QY 541 HGTSGSGISPSHSORISAGSKLSPRLQYFGOSLSGGODLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSORISAGSKLSPRLQYFGOSLSGGODLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIEMFNPREVARNVFECDQVWKGKEAGEVRVCLVHVKSTRDLRREQIOISVVT 660
DB 617 PVLVRKAIEMFNPREVARNVFECDQVWKGKEAGEVRVCLVHVKSTRDLRREQIOISVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNSRRQTQVLGTQTCETLKLQPNICIEDPVSFVILRLNF 720
DB 677 YDLALDSGRPHSRAVFNETKNSRRQTQVLGTQTCETLKLQPNICIEDPVSFVILRLNF 736
QY 721 SLVGTPLSARGNLRPVLAEDAQLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVG 780
DB 737 SLVGTPLSARGNLRPVLAEDAQLFTALPFPEKNCNDNICODDLSITFSFMSLDCLVG 796
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPDLSYRKVSTLQNRQSRQSWRLACSSASSTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPDLSYRKVSTLQNRQSRQSWRLACSSASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVLVPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVLVPV 976
QY 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORIQCDIP 1020
DB 977 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORIQCDIP 1036
QY 1021 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
DB 1037 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1096
QY 1081 KVEPFEVNPPLPIVSGVGLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1137
DB 1097 KVEPFEVNPPLPIVSGVGLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGRVVVGAPQETVAANQORGLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGRVVVGAPQETVAANQORGLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVYKGLCFLFGSLNRQQPQK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVYKGLCFLFGSLNRQQPQK 136

QY 121 FPEALRGCPQEDSDIAPLIDSGSIIPHDFFRMKELVSTIMEQLKSKTFLSLQYSEEP 180
DB 137 FPEALRGCPQEDSDIAPLIDSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLQYSEEP 196

QY 181 RIHFTFKFQNNPNRSLIKPTQLGLGTHATGLRKVVRELFNITNGARKNAFKILPLL 240
DB 197 RIHFTFKFQNNPNRSLIKPTQLGLGTHATGLRKVVRELFNITNGARKNAFKILVVI 256

QY 241 TDGEKFGDPLGVEDVIPSLEDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVFQAN 300
DB 257 TDGEKFGDPLGVEDVIPSLEDREGVIRYVLGFGDAFRSEKSRQELNTIASKPRDHVFQV 316

QY 301 NFEALKTQNLQREKIFAIEGTQTGSSSFHEMSQEGFSAATISNGPLLTSTVGSYDWAG 360
DB 317 NFEALKTQNLQREKIFAIEGTQTGSSSFHEMSQEGFSAATISNGPLLTSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGGQSVCP 480
DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGGQSVCP 496

QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAYVLF 540
DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAYVLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTGAGQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTGAGQGHVLLRSQ 616

QY 601 PVLRVKAIMFENPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTDRRLREGQIQSVVT 660
DB 617 PVLRVKAIMFENPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTDRRLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTCTETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTCTETLKLQLPNCIEDPVSPIVLRNF 736

QY 721 SLVGTPLSAFGNLRPVLAEDAQRALTALFFPKXKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGNLRPVLAEDAQRALTALFFPKXKNCNDNICQDDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLSTRKVKSTLQNRQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPDLSTRKVKSTLQNRQSRQSWRLACESASSTEV 856

QY 841 SGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
DB 857 SGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916

QY 901 QLELPVKYAVYVMTSHGVSTKYVLTNFTASENTSRVMQHYQVSNLQGRSLPISLVFLVPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYVLTNFTASENTSRVMQHYQVSNLQGRSLPISLVFLVPV 976

QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVYVNCIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVYVNCIAVCQRIQCDIP 1036

QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1096

QY 1081 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMEGGPPGAEPQ 1137
DB 1097 KVEPFEPVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMEGGPPGAEPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-193-043-3									
Query Match 99.3%; Score 5839; DB 3; Length 1153;									
Best Local Similarity 98.9%; Pred. No. 0;									
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI	76						
Qy	61	RLQVPVEAVNMVLSGLSLAATTSPPQLACGPTVHQTCSNTYVKGKLCFLFGSNLRQOPQK	120						
Db	77	RLQVPVEAVNMVLSGLSLAATTSPPQLACGPTVHQTCSNTYVKGKLCFLFGSNLRQOPQK	136						
Qy	121	PPEARLRCPCQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTILFSLMOYSEEF	180						
Db	137	PPEARLRCPCQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTILFSLMOYSEEF	196						
Qy	181	RIHFTFKFQNNPRSLIKPITQLGRTHATGLRKVVRELFINITNGARKNAFKILFL	240						
Db	197	RIHFTFKFQNNPRSLIKPITQLGRTHATGLRKVVRELFINITNGARKNAFKILFL	256						
Qy	241	TGGEKFGDPLGYEDVPEADREGVIRYVIGVGDFAFRSEKSRQELNTVASKPRDHVQAN	300						
Db	257	TGGEKFGDPLGYEDVPEADREGVIRYVIGVGDFAFRSEKSRQELNTVASKPRDHVQAN	316						
Qy	301	NFEALKTQVQLREKIPIAIEGTQTGSSSSFEHMSQEGFSAATNSGPNLLSTVGSYDWAG	360						
RESULT 9									
US-09-688-307A-3									
; Sequence 3, Application US/09688307A									
; Patent No. 6432404									
; GENERAL INFORMATION:									
; APPLICANT: Gallatin, Michael W.									
; TITLE OF INVENTION: No. 6432404el Human Beta-2									
; FILE REFERENCE: 27866/36646									
; CURRENT APPLICATION NUMBER: US/09/688,307A									
; PRIOR FILING DATE: 2000-10-13									
; PRIOR APPLICATION NUMBER: 09/193,043									
; PRIOR FILING DATE: 1998-11-16									
; PRIOR APPLICATION NUMBER: 08/605,672									
; PRIOR FILING DATE: 1996-02-22									
; PRIOR APPLICATION NUMBER: 08/173,497									
; PRIOR FILING DATE: 1993-12-23									
; PRIOR APPLICATION NUMBER: 08/286,889									
; PRIOR FILING DATE: 1994-08-05									
; PRIOR APPLICATION NUMBER: 08/362,652									
; PRIOR FILING DATE: 1994-12-21									
; PRIOR APPLICATION NUMBER: 08/943,363									
; PRIOR FILING DATE: 1997-10-03									
; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-688-307A-3									
Query Match 99.3%; Score 5839; DB 4; Length 1153;									
Best Local Similarity 98.9%; Pred. No. 0;									
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI	76						
Qy	61	RLQVPVEAVNMVLSGLSLAATTSPPQLACGPTVHQTCSNTYVKGKLCFLFGSNLRQOPQK	120						
Db	77	RLQVPVEAVNMVLSGLSLAATTSPPQLACGPTVHQTCSNTYVKGKLCFLFGSNLRQOPQK	136						
Qy	121	PPEARLRCPCQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTILFSLMOYSEEF	180						
Db	137	PPEARLRCPCQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTILFSLMOYSEEF	196						
Qy	181	RIHFTFKFQNNPRSLIKPITQLGRTHATGLRKVVRELFINITNGARKNAFKILFL	240						
Db	197	RIHFTFKFQNNPRSLIKPITQLGRTHATGLRKVVRELFINITNGARKNAFKILFL	256						
Qy	241	TGGEKFGDPLGYEDVPEADREGVIRYVIGVGDFAFRSEKSRQELNTVASKPRDHVQAN	300						
Db	257	TGGEKFGDPLGYEDVPEADREGVIRYVIGVGDFAFRSEKSRQELNTVASKPRDHVQAN	316						
Qy	301	NFEALKTQVQLREKIPIAIEGTQTGSSSSFEHMSQEGFSAATNSGPNLLSTVGSYDWAG	360						

Db 317 NFEALKTIONQREKIPAEIGTQTGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDMAG 376
Qy 361 GVFLYTSKESKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVLCAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKESKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVLCAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
Qy 481 PRGORARWOCDAVLYGEOGQWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLYGEOGQWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFCNDQVWKGKAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFCNDQVWKGKAGEVRVCLHVOKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGPHRAVENETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGPHRAVENETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALPFPFKNCNDNICODDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLFTALPFPFKNCNDNICODDLSITFSFMSLCLVVG 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTCSINHPIPPENSEVTNITFPDVSASLGKGLKLLKANVTSENMPRTNKTET 900
Db 857 SGALKSTCSINHPIPPENSEVTNITFPDVSASLGKGLKLLKANVTSENMPRTNKTET 916
Qy 901 QLELPVKYAVVMTSHGVSSTKYNLFTASNTSRVMOHOYQVSNLQORSPLISLFLVLPV 960
Db 917 QLELPVKYAVVMTSHGVSSTKYNLFTASNTSRVMOHOYQVSNLQORSPLISLFLVLPV 976
Qy 961 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQC 1020
Db 977 RLNTQVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQC 1036
Qy 1021 FFGIQBEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSFTLLPQCGAFVRSQ 1080
Db 1037 FFGIQBEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSFTLLPQCGAFVRSQ 1096
Qy 1081 KVEPPEVPNPLPIVGVSSVGGLLLALITAAALYKLGFFKROKDMWSEGGPPGABPO 1137
Db 1097 KVEPPEVPNPLPIVGVSSVGGLLLALITAAALYKLGFFKROKDMWSEGGPPGABPO 1153

RESULT 10

US-08-476-062A-43

Sequence 43, Application US/08476062A

Patent No. 5877275

GENERAL INFORMATION:

APPLICANT: Arnaout, M. Amin

TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY

TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-476-062A-43

Query Match 98.8%; Score 5808.5; DB 2; Length 1152;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQTSNTYVVKGLCFGLFNSLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQTSNTYVVKGLCFGLFNSLRQPOK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMMKEFVSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVIPEDREGVIRYVYVGLFGDAPFRSEKSRQELNTVASKPPRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVIPEDREGVIRYVYVGLFGDAPFRSEKSRQELNTVASKPPRDHVFOAN 316
Qy 301 NFEALKTVONOLREKIPIAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDMAG 360
Db 317 NFEALKTVONOLREKIPIAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDMAG 376
Qy 361 GVFLYTSKESKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVLCAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKESKSTFINNTRVDSMDNDAYLGAAAILLRNRVQSLVLCAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
Qy 481 PRGORARWOCDAVLYGEOGQWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLYGEOGQWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 600

Db 556 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 615
Qy 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 616 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAVNETKNSRRQTVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 676 YDLALDSGRPHSRVAVNETKNSRRQTVGLTQTCETLKLQLPNCIEDPVPVILRLNF 735
Qy 721 SLVGTPLSAFNGLRPVLAEADQLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFNGLRPVLAEADQLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 915
Qy 901 QLELPVKYAVTVMTVSHGVSTKYLNFTASENTRVMQHQYQVSNLQSRPLISLVFLVPV 960
Db 916 QLELPVKYAVTVMTVSHGVSTKYLNFTASENTRVMQHQYQVSNLQSRPLISLVFLVPV 975
Qy 961 RLNOTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1020
Db 976 RLNOTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1035
Qy 1021 FFGIOEBENATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTEF 1080
Db 1036 FFGIOEBENATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTEF 1095
Qy 1081 KVEFEVNPPLPLIVGSSVGLLILALITAAALYKLGFFKRYKDKMSEGGPPGAEPQ 1137
Db 1096 KVEFEVNPPLPLIVGSSVGLLILALITAAALYKLGFFKRYKDKMSEGGPPGAEPQ 1152

RESULT 11

PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906

TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US96-01314-43

Query Match 98.8%; Score 5808.5; DB 5; Length 1152;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFFQENARFGQSVVQLQGSRRVVVGAPOEIVAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFFQENARFGQSVVQLQGSRRVVVGAPOEIVAANQRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHTQCSNTYVYKGLCFPLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHTQCSNTYVYKGLCFPLFGSNLRQPOK 136
Qy 121 FPEALRGCCPOEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKSKTLFSLMQVSEEF 180
Db 137 FPEALRGCCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOLKSKSKTLFSLMQVSEEF 196
Qy 181 RIHFTFEKFNQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFEKFNQNNPNRSLVKPITQLLGRTHRTATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVTPELDREGVIRYVLGFGDAFRSEKSRQBELNTVASKPPRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVTPELDREGVIRYVLGFGDAFRSEKSRQBELNTVASKPPRDHVFOAN 316
Qy 301 NFEALKTQNLQREKIPALEGTOTGSSSFHEHMSQEGFSAAITSNGLPLSTVGSYDWAQ 360
Db 317 NFEALKTQNLQREKIPALEGTOTGSSSFHEHMSQEGFSAAITSNGLPLSTVGSYDWAQ 376
Qy 361 GVFLYTSKEKSTFINMTRVDSNMNDAYLGAAAAIILNRNVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSNMNDAYLGAAAAIILNRNVQSLVLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTRGGQSVCPPL 480
Db 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEQTRGGQSVCPPL 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAYVLF 540
Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPEGEDNRGAYVLF 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 600
Db 556 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 615
Qy 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 616 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAVNETKNSRRQTVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 676 YDLALDSGRPHSRVAVNETKNSRRQTVGLTQTCETLKLQLPNCIEDPVPVILRLNF 735
Qy 721 SLVGTPLSAFNGLRPVLAEADQLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFNGLRPVLAEADQLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 915

[illegible]

Db	497	PRG - RARWQCDVLYGEEQQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEENRGAUYLP	555
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFQGSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ	600
Db	556	HGTSGSGISPSHSQRIAGSKLSPRLQYFQGSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ	615
Qy	601	PVLRVKAIEMFNPREVARNVFCNDQVVGKAGEVRVCLHVQKSTRDRRLRGQIQSVVT	660
Db	616	PVLRVKAIEMFNPREVARNVFCNDQVVGKAGEVRVCLHVQKSTRDRRLRGQIQSVVT	675
Qy	661	YDLALDSGRPHSRVAFNETKNSRTRTOVLGLTQTCTETLKQLPNCIEDPVSPVILRLNF	720
Db	676	YDLALDSGRPHSRVAFNETKNSRTRTOVLGLTQTCTETLKQLPNCIEDPVSPVILRLNF	735
Qy	721	SLVGTPLSAFNLRPVLAEDAQRLEFTALFFFEKNCNDNICODDLSITFSFMSLDCLVVG	780
Db	736	SLVGTPLSAFNLRPVLAEDAQRLEFTALFFFEKNCNDNICODDLSITFSFMSLDCLVVG	795
Qy	781	GPREFVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV	840
Db	796	GPRESVTVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV	855
Qy	841	SGALKSTSCSINHPIPEPENSEVFNITFDVDSKASLGNKLLKANVTSNNMPRTNKTBF	900
Db	856	SGALKSTSCSINHPIPEPENSEVFNITFDVDSKASLGNKLLKANVTSNNMPRTNKTBF	915
Qy	901	QLELPVKYAVYVTVTSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSPLISLVFLVPV	960
Db	916	QLELPVKYAVYVTVTSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSPLISLVFLVPV	975
Qy	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCIAVCQRIQCDIP	1020
Db	976	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCIAVCQRIQCDIP	1035
Qy	1021	FFGQIEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPQOGAFVRSQTET	1080
Db	1036	FFGQIEEFNATLKGNLSFDWYIKTSHNHLLIVSTAEILFNDSVFTLLPQOGAFVRSQTET	1095
Qy	1081	KVBPFFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKXDMMSGGPPGAEQ	1137
Db	1096	KVBPFFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKXDMMSGGPPGAEQ	1152
RESULT 13			
US-08-476-062A-44			
: Sequence 44A, Application US/08476062A			
: Patent No. 5877275			
: GENERAL INFORMATION:			
: APPLICANT: Arnaout, M. Amin			
: TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY			
: TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS			
: NUMBER OF SEQUENCES: 53			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Fish & Richardson P.C.			
: STREET: 225 Franklin Street			
: CITY: Boston			
: STATE: MA			
: COUNTRY: US			
: ZIP: 02110-2804			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Diskette			
: COMPUTER: IBM Compatible			
: OPERATING SYSTEM: Windows95			
: SOFTWARE: FastSeq for Windows Version 2.0			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/476,062A			
: FILING DATE: 07-JUN-1995			
: PRIOR APPLICATION DATA:			
: APPLICATION NUMBER: 08/216,081			
: FILING DATE: 21-MAR-1994			
: APPLICATION NUMBER: 07/637,830			
: FILING DATE: 04-JAN-1991			
: APPLICATION NUMBER: 07/539,842			

FILING DATE:	18-JUN-1990
APPLICATION NUMBER:	07/212,573
FILING DATE:	28-JUN-1988
ATTORNEY/AGENT INFORMATION:	
NAME:	Freeman, John W.
REGISTRATION NUMBER:	29,066
REFERENCE/DOCKET NUMBER:	00786/068003
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	617/542-5070
TELEFAX:	617/542-8906
TELEX:	200154
INFORMATION FOR SEQ ID NO:	44:
SEQUENCE CHARACTERISTICS:	
LENGTH:	1163 amino acids
TYPE:	amino acid
TOPOLOGY:	linear
MOLECULE TYPE:	protein
US-08-476-062A-44	

Query Match	58.6%; Score 3446; DB 2; Length 1163;
Best Local Similarity	60.5%; Pred. No. 4.9e-278;
Matches	683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

Qy	1	FNLDTENAMTFOENARGFGOSVQLGSRVVVGAPQEIVAAANRGSLYQCDYSTGSGCEPI	60
Dd	20	FNLDEBELTAFRVDSAGFGDSVQVYANSWVVGAPQKITAAANTGGLYQCQGYTGACEPI	79
		: : : : : : : : : : :	
Qy	61	RLOVPVEAVNMSLGLSLAATTPPOLLACAGPTVHQTCSENTYYVKGCLCFLFGSNLRQQOPK	120
Dd	80	GLQVPPEAVNMSLGLSLASTTSPOLLACAGPTVHECGRNMYLTGLCFLLGPT--QLTOR	137
		: : : : : : : : : :	
Qy	121	FPEALRCGPQBDSDFIALIDGSGSIIIPHFMRMKELVSTIMEQLKKSKTLFSLMQYGEFF	180
Dd	138	LPVSREQCPREQDIFVLIDGSGSISSRFATMNFRAVISQFORPSTQSFSLMQFSNKPF	197
		: : : : : : : : : :	
Qy	181	RIHTTFEFQNPNRSLIKPIITOLLGRTHATGLRKVVRELFNITNGARKNAKILFL	240
Dd	198	QTHFTFEFRRTSNPLSLLASVHOLQGTYTATAIQNVVHLPHASYGAREDDATKLIVI	257
		: : : : : : : : : :	
Qy	241	TDGKFGDGLGVEDVIPELDREGVRVYVLGFDAFRSEKSQEOELNTVASKEPRDHVFQAN	300
Dd	258	TDGKKGESLDYKDVI PWADAAGIIRVAIGVLAFQNRNSWKELNDIAKSPSQEHI FKVE	317
		: : : : : : : : : :	
Qy	301	NFEALKTVQNOLREKIFAIEGTQTGSSSSFHEMSQEGFSAAITSGPLLSTVGSYDWAG	360
Dd	318	DFDALKDIOQLQEKI PAIEGTETTTSSSSFELEMAQEGFSAVFTPDGPVLGAVGSTWSG	377
		: : : : : : : : : :	
Qy	361	GVLFLTYSKEKSTFTNMTRVDSMDADYLGYAAAIIILNRVQSULVGAPRYOHIGLVAMFR	420
Dd	378	GAFLYPPNMSPTFINMSQENVDMRDLSYGSTELAKMWGVQSVLGAAPRYOHTGKAIFT	437
		: : : : : : : : : :	
Qy	421	QNTGMWESNANVKGTQICAYFGASICSVDVDSNGSTDVLVLGAPHYYEQTRGGQVSCVPL	480
Dd	438	QVSRQWRMAEVGTGTQIGSYFGASLCUSDVDTDGSTDLVLGAPHYYEQTRGGQVSCVPL	497
		: : : : : : : : : :	
Qy	481	PRGORARWCQDAVLYGEGQPWRFGAALTVLGDVNGDKLTDVAITGAPGEEDNRGAVYLF	540
Dd	498	PRGNR-RWMCDAVLYGEGHPWRFGAALTVLGDVNGDKLTDVVI GAPGEENRGAVYLF	556
		: : : : : : : : : :	
Qy	541	HGTSGSGISPSHSQR IAGSKLSPRLQYFGSLSGGQDLTMGLVDLTVGAQHVVLLRSQ	600
Dd	557	HGVLGPSISP SHSQR IAGSQLSRRLQYFCQALLSGGQDLTDQGLVDLAVGARGQVLLLRTR	616
		: : : : : : : : : :	
Qy	601	PVLRVKATMENPREVANFECDNDVVYKGEAGEVRVCLHVOKSTRDLRLEGQIQSVVT	660
Dd	617	PVLWMVGSMQPIPAIEIPSAFECREQVVSEOTLVQSNICLYIDKRKNLLGSRDQSSVT	676
		: : : : : : : : : :	
Qy	661	YDALDSCRGRHSRAVFNETKNSRQTQVIGLTQTCETLKQLPNCIEDPVSPVILRLNF	720
Dd	677	LDLALDPGLSPRATFQETKNSLSRVRLGLKAHCENFNLLPSCVEDSVTPI TLRLNF	736
		: : : : : : : : : :	
Qy	721	SLVGTPTL\$AFGNLRPVLAEADAQRLTALPPEFKNCGNDNI CODDSLITTSFMSMLDCLVVG	780
		: : : : : : : : : :	

QY 1 FNLDTENAMTFOENARGFGOSVVOQLGSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB 20 FNLDTEELTAFAFRVDSAGFSDVQVYANSWVVGAPQKITAANQTGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVVKGLCLFLGSLNLRQOPK 120
DB 80 GLQVPPEAVNMSLGLSLASTTSPQQLACGPTVHHECGRNMYLTGLCFLGPT--QLTOR 137
QY 121 FPEARLGCQBEDSDIAFLIDGSGSIIPHDPRMKELVSTIMBOLKSKTLFSLMOYSEEF 180
DB 138 LPVSRQECPRQODIVFLIDGSGSISSRNFAFMNFRAVISOQFORPSTQFSLMQFSNKF 197
QY 181 RHFTFKFQNNPNRSLIKPTTOLLGRTHATGLRKVVRELFTNMGARKNAKFLIFLL 240
DB 198 QTHFTFEFRRTSNPLSLASVHQLGQFTYTATAIQNVRHLFHASYGARRDAIKILIVI 257
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAFRSEKSRQELATVASKPRDHVFOAN 300
DB 258 TDGKKEGSDLYDKVDPIMADAAGIIRYAGVGLAFQNRNWKELNDIASKPSOEHIKVE 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDWAG 360
DB 318 DFDALDKIQNLKEKIFAIEGTETSSSFELEMAQEGFSAVFTPDGVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTPINNTRVDSMDNDAYLGAAAILNRNVSQSLVGLGAPRYQHIGLVAMPR 420
DB 378 GAFLYPPNMSPTPINNSQENVDNRDSYLGSTELALWKGVSQSLVGLGAPRYQHTGRAVFT 437
QY 421 QNTGMWESNANVKGTOIGAFYGASLCSVDVDSNGSDTLVILGAPHYETRGQVSVCLP 480
DB 438 QVSRQRMKAETVGTQIGSYFGASLCSVDVDGSTDVLVILGAPHYETRGQVSVCLP 497
QY 481 PRGQARMOCDAVLYEQOQPMGRFGAALTALVDVNGDKLTDVAIGAPGEEDNRGAVILF 540
DB 498 PRGWR-RWCDVLYEQOHPWGRFGAALTALVDVNGDKLTDVIGAPGEENRGAVILF 556
QY 541 HGTSGGSIIPSHSQRIAGSKLSPRLOQYFQOSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
DB 557 HGVLGPSISPSHSQRIAGSKLSPRLOQYFQOALSGGQDLTODGLVLDVAVGARGOVLLRTR 616
QY 601 PVLRLVKAIMEFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGOQSUVT 660
DB 617 PVLWGVSMQFIPAEIPRFAFECREQVSEQTLVQSNICLYDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
DB 677 LDALDPGLSPRAFTQETNRSLSRVVLGLKHCENFNLLPSCVEDSVTPITRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFPALFPFKNCGNDNICQDDLSITPSFMSLDCLVVG 780
DB 737 TLVGKPLLAFLNRPMLAALAQRYFTASLPFEKNCGADHICQDNLGISPSFPLKSLLVG 796
QY 781 GRREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSRWSFLACESASSTEV 840
DB 797 SNLELNAEVMVWMDGEDSVTITFTSHAPAGLSYRYVAEQKQQLRSLHLTCDSPAVG-- 854
QY 841 SGALKSTSCSINHPIPPENSEVFTNTPDVSQKSLGNKLLKANVTSENMMPTNKTEF 900
DB 855 SGTWSTSCRIINHLIPRGAQOITFLATFDVSPKAVLGDRLLLTANVSSENNTPRISKTF 914
QY 901 QLELPVKYAVYVMTSHGYSTKYLNFTAS-ENTSRVMQHYQVSNLQGRSLPISLVFLVP 959
DB 915 QLELPVKYAVYVMTSHGYSTKYLNFTAS-ENTSRVMQHYQVSNLQGRSLPISLVFLVP 974
QY 960 VRLNQTVWDRQVTFSENLSSTCHTKERLPHSDPLAELRKAIVVNSIAVCQRIQCDI 1019
DB 975 VELNQCAVMDVEVSHPNQPSLRCSSEKIAFPASDPLAHIQNPVLDSCSIAGLCFRCDV 1034
QY 1020 PFFGQEBENATLKGNSLFDWIKTSHNHLIVSTAEILFNDISVFTLLPGQCAFVRSOTE 1079
DB 1035 PFSVQEBELDTLKGNSLFDWIKTSHNHLIVSTAEILFNDISVFTLLPGQCAFVRSOTE 1094
QY 1080 TKVEPPEVNPPLPLIVSSVGGLLALLITALIYALYKLGFFKQYKQMDMSE 1128

DB 1095 TVLEKYKVNPTPLIVGSSIGGLLLALLITAVLYKVGFQKQYKEMEE 1143
RESULT 15
US-08-173-497-4
; Sequence 4, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-4
Query Match 58.2%; Score 3423; DB 1; Length 1163;
Best Local Similarity 60.2%; Pred. No. 4.1e-276;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFOENARGFGOSVVOQLGSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB 20 FNLDTEELTAFAFRVDSAGFSDVQVYANSWVVGAPQKITAANQTGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVVKGLCLFLGSLNLRQOPK 120
DB 80 GLQVPPEAVNMSLGLSLASTTSPQQLACGPTVHHECGRNMYLTGLCFLGPT--QLTOR 137
QY 121 FPEARLGCQBEDSDIAFLIDGSGSIIPHDPRMKELVSTIMBOLKSKTLFSLMOYSEEF 180
DB 138 LPVSRQECPRQODIVFLIDGSGSISSRNFAFMNFRAVISOQFORPSTQFSLMQFSNKF 197
QY 181 RHFTFKFQNNPNRSLIKPTTOLLGRTHATGLRKVVRELFTNMGARKNAKFLIFLL 240
DB 198 QTHFTFEFRRTSNPLSLASVHQLGQFTYTATAIQNVRHLFHASYGARRDAIKILIVI 257
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAFRSEKSRQELATVASKPRDHVFOAN 300
DB 258 TDGKKEGSDLYDKVDPIMADAAGIIRYAGVGLAFQNRNWKELNDIASKPSOEHIKVE 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDWAG 360

Db 318 DFDALDKIQNLKEKIFAIEGTETISSSSFELEMAQEGSAVFTPDGPVLGAVGSFTWSG 377
Qy 361 GVFLYTSKEKSTFINMTRVDSDDMDAYLGYAAAIIILNRVQSLVGLGAPRYOHICGLVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVDMDSYLSTELALWKGVQSLVGLGAPRYOHIGKAVIFI 437
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQSVSCPL 480
Db 438 QVSRQWEMKAEVITQIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEQTGGQSVSCPL 497
Qy 481 PRGORARQCDAVLVYGGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYLIF 540
Db 498 PRGWR-RWMCDAVLVYGGQGPWGRFGAALTVLGVDVNGDKLTDVVGAPGEEENRGAYLIF 556
Qy 541 HGTSGSGISPSHSORISAGSLSPLOYFGQSLSGQDITMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLGPSISPSHSORISAGSLSPLOYFGQSLSGQDITMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFENPREVARNVECNQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSWIT 660
Db 617 PVLWGVSMQIPAEIPRSAFECEQVVEQTLVQSNICLYIDKRSKNLGSRLQSSVT 676
Qy 661 YDLALDGRPHSRFAVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 LDLALAPGRLSPRAIFQETKNSLRVRVLGKAHCENFNLLPSCVEDSVIPIILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAORLFTALFPFEKNCNDNICODDLISITFSFMSLDCLVWG 780
Db 737 TLVGKPLLAFLNRLPMLAALAQRYFTASLPFEKNCAGADHICQDNLGIFSFPGLKSLVG 796
Qy 781 GPREFNTVTVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRORSORSWRLACESASSTEV 840
Db 797 SNLELNAEVMWVNDGSDSYGTTITFSPAGLSYRYVAEGQKQGOIRSLHLC--CSAPVG 854
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 855 SGTWSTSCRINHILIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914
Qy 901 QLELPVKYAVYVWVTSKYNLFTAS-ENTSRVMOHQYOVNSLQORSPLPSLVFLVP 959
Db 915 QLELPVKYAVYVWVSSHQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSINFMPV 974
Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVMNCIAVCQIOCDI 1019
Db 975 VELNQEAVWVDEVSHPONPSLRCSSEKIAPPASDFLAHQNPVLDCSIAGCLRFRCDV 1034
Qy 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQTE 1079
Db 1035 PSFSVQEELDFTLKGNLSFGWVROILQKKSVSVSAEIIFDTSVYSQLPQGEAFMRAQTI 1094
Qy 1080 TKVEPFEVNPPLPIVGVSSVGLLLALITALIYKLGFFKQYKDMHSE 1128
Db 1095 TVLEKYKVHNPILPIVGVSSIGGLLLALITALIYKLVGVFFKQYKEMMEE 1143

Search completed: November 25, 2003, 14:23:13
Job time : 17.5618 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 14.9579 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFCQ.....FKROYKDMSEGPGAEFQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1	RWHULB
2	4460	75.8	1133	2	S00551
3	3469	59.0	1153	1	RWHULC
4	1543.5	26.2	1170	2	S03308
5	1527.5	26.0	1163	2	I56126
6	1140	19.4	1179	2	A53213
7	1093.5	18.6	1151	2	A45226
8	1071	18.2	1170	2	I45314
9	1057	18.0	1178	2	S44142
10	1055	17.9	1180	2	A35854
11	1054	17.9	1181	2	A33998
12	663	11.3	1039	2	A41131
13	642	10.9	1038	2	S06046
14	630	10.7	1035	2	I58409
15	614.5	10.4	1041	2	T31437
16	579.5	9.8	1054	2	JC7294
17	572.5	9.7	1051	2	A35761
18	567.5	9.6	1053	2	I55534
19	555.5	9.4	1033	2	S44250
20	546.5	9.3	1034	2	A36108
21	535	9.1	1044	2	T10050
22	534.5	9.1	1072	2	A38457
23	532	9.0	1049	2	A27079
24	532	9.0	1073	2	B36429
25	530.5	9.0	1051	2	A40021
26	526	8.9	1048	2	A27421
27	525.5	8.9	1091	2	A41543
28	517	8.8	1044	2	S16516
29	496.5	8.4	1146	2	S40311

ALIGNMENTS

RESULT 1

RWHULB

cell surface glycoprotein CD11b precursor [validated] - human

N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A45526; A26091; I52567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3,
B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:Cross-references: GB:J03925; NID:gl87284; PIDN:AAA59544.1; PID:G307148

A:Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor
A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499,501-965 'P', 967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAA-

A:Note: the authors translated the codon TAC for residue 1129 as Thr

A:Note: part of this sequence, including the amino end of the mature protein, was conf

R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally
A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:gl80018; PIDN:AAA58410.1; PID:G553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe
A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept
A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:Cross-references: GB:J04145; NID:gl89068; PIDN:AAA59903.1; PID:G386975

integrin - fruit f

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526; MUID:93123748; PMID:8419480
A:Accession: A46526
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499, 501-1153 <FL>
A:Cross-references: GB:S52227; NID:9263047; PIDN:AB24821.1; PID:9263049
A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26091
A:Molecule type: protein
A:Residues: 17-31 <PI>
A:Experimental source: granulocytes
R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A:Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: 152567; MUID:92144986; PMID:1346576
A:Accession: 152567
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:Cross-references: GB:M84477; NID:g180184; PIDN:AA51960.1; PID:g553219
C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genetics:
A:Gene: GDB:ITGAM; CR3A
A:Cross-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F:17-1108/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <WAA>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240, 391, 469, 693, 977, 735, 802, 881, 901, 912, 941, 947, 979, 994, 1022, 1045, 1051, 1076/Binding

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTQENARGFGQSVVQLQSGSRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTQENARGFGQSVVQLQSGSRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI	76
Qy	61	RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVKGCLCFLGSNLRQOQPK	120
Db	77	RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVKGCLCFLGSNLRQOQPK	136
Qy	121	FPEALRGCPQSDSIAFLIDGSGSIIIPHDPRMKWSTVMEQLKSKTFLSLMQYSEEF	180
Db	137	FPEALRGCPQSDSIAFLIDGSGSIIIPHDPRMKWSTVMEQLKSKTFLSLMQYSEEF	196
Qy	181	RIHFTFKFQNNPNRSIIKPIITQLLGRTHATGLRKVRELFINITNGARKNAFKILPLL	240
Db	197	RIHFTFKFQNNPNRSIIKPIITQLLGRTHATGLRKVRELFINITNGARKNAFKILVVI	256
Qy	241	TDGEKFGDPLGVEDVPELDREGVIRYVIGVDAPRSEKSKOELNTVASKPPDRHVFOIN	300
Db	257	TDGEKFGDPLGVEDVPELDREGVIRYVIGVDAPRSEKSKOELNTVASKPPDRHVFOVN	316

Qy	301	NFEALXTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAALTSNGPLLLSTVGSYDWAG	360
Db	317	NFEALXTIQNQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAALTSNGPLLLSTVGSYDWAG	376
Qy	361	GVFLYTSKSKSTFINNTRVDSQNDMDAYLGAALILNRNVQSLVLCAPRQVHIGLVAMPR	420
Db	377	GVFLYTSKSKSTFINNTRVDSQNDMDAYLGAALILNRNVQSLVLCAPRQVHIGLVAMPR	436
Qy	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGRGQVVCPL	480
Db	437	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGRGQVVCPL	496
Qy	481	PRQQRARWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAICAPGEEDNRGAVILF	540
Db	497	PRQQRARWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAICAPGEEDNRGAVILF	556
Qy	541	HGTSGSGISPHSHSORIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ	600
Db	557	HGTSGSGISPHSHSORIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ	616
Qy	601	PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660
Db	617	PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	676
Qy	661	YDLALDSGRPHSRAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF	720
Db	677	YDLALDSGRPHSRAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF	736
Qy	721	SLVGTPLSAFGNLRPVLAEADAQRLFTALPPFEKNCGNDNICDDLSITFSFMSLDCLVVG	780
Db	737	SLVGTPLSAFGNLRPVLAEADAQRLFTALPPFEKNCGNDNICDDLSITFSFMSLDCLVVG	796
Qy	781	GPREFNVTVTRNDEGDSYRTQVTFPPDLDSYKRVSTLQNRQSQRSLACSSASSTEV	840
Db	797	GPREFNVTVTRNDEGDSYRTQVTFPPDLDSYKRVSTLQNRQSQRSLACSSASSTEV	856
Qy	841	SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANTVTSENNMPTNKTEF	900
Db	857	SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANTVTSENNMPTNKTEF	916
Qy	901	QLSELPVKYAVYVMVTSHGVSSTKYLNFTASNTSRVMOHQVQVSNLQORSIPISLVFLPV	960
Db	917	QLSELPVKYAVYVMVTSHGVSSTKYLNFTASNTSRVMOHQVQVSNLQORSIPISLVFLPV	976
Qy	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1020
Db	977	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1036
Qy	1021	FFGIQBEFNATLKGNSLFDWYIKTSNNHLLIVSTAEILFNDSVFTLLPGQGAFFVRSQTE	1080
Db	1037	FFGIQBEFNATLKGNSLFDWYIKTSNNHLLIVSTAEILFNDSVFTLLPGQGAFFVRSQTE	1096
Qy	1081	KVEPPEFVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ	1137
Db	1097	KVEPPEFVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ	1153

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N:Alternate names: complement-3 receptor alpha chain
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999
C:Accession: S00551; I59078
R:Pytela, R.
EMBO J. 7, 1371-1378, 1988
A:Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with th
A:Reference number: S00551; MUID:88312584; PMID:3044779
A:Accession: S00551
A:Molecule type: DNA
A:Residues: 1-1153 <PYT>
A:Cross-references: EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52983
A>Note: the authors translated the codon CAC for residue 569 as Gln

R.Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
A:Reference number: I59078; MUID:86287312; PMID:2942940
A:Accession: I59078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Cross-references: GB:M14293; NID:q198993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 75.8%; Score 4460; DB 2; Length 1153;
Best Local Similarity 73.8%; Pred. No. 2.2e-303;
Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;

Qy 1 FNLDTENAMTFQENARGFQSVVQJGSRVVVGAQOEIVAAANQRSLYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFQENAKFGQNVVQLGGTSVVVAAPQEAQAVNQTGALYQCDYSTGRCHI 76

Qy 61 RLQVPVEAVNMSIGLSLAATSPPOLACGPVTHQTCSENTYVKGICFLFGSNLRQQPK 120
Db 77 PLQVPPEAVNMSIGLSLAATSPPOLACGPVTHQNCENTYVNGLCYLFGSNLRPPQQ 136

Qy 121 FPEALRGCPQEDSDTAFLIDSGSIIPHDFRMKEWSTVMEQLKSKTFLSLMQYSBEF 180
Db 137 FPEALRECPQESDIVFLIDSGSINNIDFQMKKEFVSTVMEQFKSKTFLSLMQYSDEF 196

Qy 181 RIHFTKFEQNNPRSLIKPTQLGLRTHATGLRKVVVRELFINITNGARKNAKIFEL 240
Db 197 RIHFTFNDPKRPSRSHVSPKQLNGRTKTASGIRKVVRELPHKTNAGARENAKILVVI 256

Qy 241 TQGEKFGDPLGYEDVITPELDREGVIRYVIGVGDAPRSEKSRQELNATVASKPRDHVFQIN 300
Db 257 TDGEXFGDPLDYKDIPEADRAGRVIRYVIGVGNANFNKQSRRELDTITASKPAGEHVQVD 316

Qy 301 NFEALKTIQNLQREKIFALEGTQTSSEHSEMSQEGFSAITNGPLLSVTGVSVDWAG 360
Db 317 NFEALNTIQNLQEKIFALEGTQTSSEHSEMSQEGFSAITNGPLLSVGSVDFWAG 376

Qy 361 GVFLYTSKESFTFNNTRVDSNDNDAYLGYAAAILRNVRQSLVLGAPRYQHIGLVAMPR 420
Db 377 GAFLYTSKDKVTFINTTRVDSNDNDAYLGYASAVILRNVRQSLVLGAPRYQHIGLVAMPR 436

Qy 421 QNTGWESNANVKGTOIGAFGASLCSVDVDSNGSTDVLICGAPHYETRCGQSVCP 480
Db 437 ENFGTWEPHTSTKSGQISGYFASLCSVDMDADGNTNLILIGAPHYETRCGQSVCP 496

Qy 481 PRGQARMQCDVAVLGEQGPWGRFGAALTLDGVNKGDLTDVAIGAPDEENRGAVYLF 540
Db 497 PRG-RARWQCEALLHGDQHPWGRFGAALTLDGVNKGDLTDVAIGAPDEENRGAVYLF 555

Qy 541 HTSGSGISPSHSQRTAGSLPRLOYFQOSLSGGQDLMGDLVLTGVAQGHVLLRSQ 600
Db 556 YGASIASLSSASHRIIGAHFSPGLQYFQOSLSGGKDLTMDGLMDLAVGAQGHVLLRAQ 615

Qy 601 PVLRVKAMFNPREVARNVFECDNVQVKGKAGEVRVCLHVOKSTRDLRGQTSVVT 660
Db 616 PVLRLAEATMEFSPKPKVARSVFACQEQVLKNKADAGEVRVCLRVKNTKDLRGDQTSVT 675

Qy 661 YDLALDSGPHRPAFVNETKSTRQTVLGTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 676 YDLALDPVRSRTRAFDETKNNTRRTQVFLGQKCEETLKLTPDCVDDSVSPVILRLNY 735

Qy 721 SLVGTPLSAFGLNLRPVLAEDAORFLTALFPFPEKNCNDNICDDLSITFSPMSLCLVVG 780
Db 736 TLVGEPLRSFGLNLRPVLAEDAORFFTFMFPPEKNCNDNICDDLSITMSANGLDVLVG 795

Qy 781 GPREFNVTVTVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRL-ACESASSTE 839
Db 796 GPQDFNMSVTLNRNDEGDSYGTQVTVYFSGLSYRKDSASQNPITKPKPWFVPAESSSE 855

Qy 840 VSGALKSTSCSINHPIIPENSEVFNITFDVDSKASLGNKLLKLLKANVTSENMPRTNTE 899
Db 856 GHGALKSTTWNINHPIIPANSEVFNITFDVDSHASEFGNKLKLLKAIIVASENNMSRTHTK 915

Qy 900 FQLELPVKYAVYVMVTSHTKYNLFTASNTSRVMOHOYQVSNLQORSPLISLVFLVP 959
Db 916 FQLELPVKYALYMLVTSDESSIRYLNFTASMTSKVIOHQYQFNLLGQRSUPSVVFWIP 975

Qy 960 VRLNQTVIWRDPQVTFSENLSSTCTKTERLPSHSDFLAELRKAPVNCISVACORIOCDI 1019
Db 976 VQINNVTWMDHPQVIFSONLSSACHTEQKSPHNSFRDQLETPVLNCSVAVCKRIQCDL 1035

Qy 1020 PFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
Db 1036 PPSNTQEIFNFTLKNLSFDWYIKTSHGHLLVSTELFNDSAFALLPGQESVYRSKTE 1095

Qy 1080 TKVSPFEVNPPLIVGVSSVGLLLALITAAALKYKGFKKQYKDMSEGGPPGAEPQ 1137
Db 1096 TKVSPYEVHNPVPLIVGVSSIGLLVLLALITAGLYKGFKKQYKDMMEAAAPQDAPPQ 1153

RESULT 3
RWHUIC
cell surface glycoprotein CD11c precursor - human
N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999
C:Accession: A36584; A35543; S00864
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A36584
A:Contents: erratum
A:Accession: A36584
A:Molecule type: DNA
A:Residues: 1-1163 <COR>
A:Note: this revision to the sequence from reference A35543 includes the carboxyl end
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule
A:Reference number: A35543; MUID:90153906; PMID:2303426
A:Accession: A35543
A:Molecule type: DNA
A:Residues: 1-834 <CO2>
A:Note: this sequence has been revised in reference A36584
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A:Reference number: S00864; MUID:86166645; PMID:3327687
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755, 'L', 757-1163 <CO3>
A:Cross-references: GB:M81695; EMBL:Y00093; NID:q487829; PIDN:AAA59180.1; PID:q487830
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119758; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat ho
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repe
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F:20-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61.89,392.697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status p

175 QYSEPRHIFPFKEFQNNPRSLIKPIITQLLGRTHATGLRKVVRELNIITNGARKNAP 234
197 QFSTSYKTEBDFSDYVWKWDPDALLKHVHMLLTNTFGAINVATEVFEELGARPDAT 256
235 KILFLITDGEKFDGPIGYEDVPELDRGVIRVIGVDAFRSEKSRQBELNTVASPPRD 294
257 KVLIIITDGE--ATDSGNTDAKD-----IIRYIIGIKGHFQTKESQETHLKFASKPASE 309
295 HVFOINNFEALTKIQNLREKIFAISGTOTGSSSSFEHEMSQSGPSAAITNSGPLLSTVG 354
310 FVKILDTFEKLDLFLTELQKIVIEGTSKQDTSFNMLSSGSIADLSRGHVVAVGAVG 369
355 SYDWAGGVF--LYTSKEKSTFINNTRVDSMDNDYLGAAA--IILNRVQSLVLAGAPRYQH 412
370 AKDWAGGLDKADLQDDTFIGNELPTPEVRAGYLGVTWLPSPRKQKTSLLASGARPRYQH 429
413 IGLVAMFR--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQT 470
430 MGRVLLFQEPQGGHWSQVTHGTQIGSYFGELGCDVDQDGETELLIGAPLFYGEQ 489
471 RGGQVSVCLPRQGRARWCDAY--LYGEOGQPMGRFGAALTVLGVDVNGDKLTDVAIGAP 528
490 RGRVFIY-----QRRQLGFEVESELOQDGPYPLGRFGEAITALTIDNGDLVDVAVGAP 544
529 GEEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTV 588
545 LEE--QGAVYIFNGRHG--GLSPQSPQRIEGTVLSGIQWFGRSIHGVKLEGGDLADVAV 601
589 GAQGHVLLRSQVPLRVKAIMENPREVARNVPECNDQV--KKEAGEVRVCLVHVKQSTR 647
602 GABSONIVLSREPVDMVIMLSPSPAEIPVHEVECSYSTSNKMKEGVNIITCFQI--KSLY 660
648 DRLREGIOQSVVTDIALDSGRPHSRVNETKNSRRQTVGLTQTCETLKLQLPNCI 707
661 PQP--QQRVANLYTTLQOHRTRRGLPFGGRHELRNIAVT--TMSCTDFSFPFVPCV 718
708 EDPVSPVLRNLNPSL---VGTPLS--AFGN-----LRPVLAEADAQLFTALPFERNCGN 757
719 QDLISPINVSLNPSLWEEGTPDQRAQKQDIPPLRPSLSHSETWEI-----PFERNCGE 773
758 DNICQDDLSITFSFMSLCLVGVGPREFNVTVRNDGSDSVRTQVTFPPFDLSVRKVS 817
774 DKKEANLRVSPARSRALRLTAFASLSVELSLSNLEEDAYVWQLDLHFPGLSPRKVE 833
818 TLQNRQSORSLRACES--ASSTEVSGALKSTSCSINHPIFPENSSEVTNITFDVDSKAS 875
834 ML---KPHSQIPVSCCELPEERLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNS 887
876 LGNKLILLKANVTSENN---MPRTNKTEFQLELPVKYAVVMVVTSHGVSTKYLNFNTASEN 931
888 WGDVSELHANVTCNEDSDLEDNSATTI---IPILYPINILLIQOEDSTLYVSTFPKGP 944
932 TRVMQHOVQV---SNLQORSLP--ISLFLVPLVRLNQTVIMDRPQVTFSENLSSTCHTK- 986
945 KIHQVKHMYQVRIQPSIHDNIPTLEAVVGVPPQPPSEGPITHQWSVQMEPPV--PCHYED 1002
987 -ERLPSHSD--FLAELRKAPVNCSTAVCQRICQDIPFGIGIEEFNATLKNLSFDWYIK 1043
1003 LERLPDAABFCUPGALFRCPW-----FRQEILVQVIGTLELVEGIE 1044
1044 TSHNHLIIVSTABILFNDVSFTLLPGQGAFFVRSQTEKVEPPEVPNPLPLIVGSSVGGLL 1103
1045 AS-SMFSLCSLSISFNSKGFHLYGSNALS-AQVVMKVDVYVEKQMLYLYVLSGIGLL 1102
1104 LIALITAALYKGFGRKQYKDMMSSEG-GPPGAP 1136
1103 LULLIIVLYKVGFFRNLKERKAGRVNGIP 1136

RESULT 5
I56126
lymphocyte

fuction-associated molecule-1-alpha - mouse

C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C:Accession: I56126
R:Kaufmann, Y.; Tseng, B.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A:Reference number: I56126; MUID:91268576; PMID:2051027
A:Accession: I56126
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residue: 1-1163 <RGS>
A:Cross-references: GB:I560778; NID:gl98785; PIDN:AAA39426.1; PID:gl98786
C:Genetic: C:
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat h
C:151-315/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match 26.0%; Score 1527.5; DB 2; Length 1163;
Best Local Similarity 34.1%; Pred. No. 3.6e-98;
Matches 398; Conservative 217; Mismatches 463; Indels 89; Gaps 36;

QY 1 FNLDTENAMTFOENA-RGFGQSVVQLQGSRRVVVGAPOEIVAAQNRGSLYQCDYSTGSCPEP 59
Db 24 YNLDRTPRTQSFLAQAGRHFGYQVLTQIDG-VVVGAPGE---GDNTGGLYHCHRTSEFCQP 79
QY 60 IRLQVPVEAVNMNLSGLSLAATTPPQLLACGPTVHTQTCSENTYVVKGLCFGLGSLNRQQPQ 119
Db 80 VSLH--GSNHTSKYLQMTLATDAKGSLLACDPELSRTCDQNTYLSGLCYLFPQSLEGPM 138
QY 120 KPEALRQCPQEDSDIAFLIDGSGSII PHDFRMKEMVSTVMEQLKSKTFLSLMOYSEE 179
Db 139 QNRPAYQECMKGVLDVFLFDGSQLDRKDPKLEFEMKDMVRKLSNTSYQAAVQFSTD 198
QY 180 PRIHTEKEF--QNNNPRSLIKPIITQLLGRTHATGLRKVVRELNIITNGARKNAPKIIF 238
Db 199 CTEETFIDYVQKQNPVLLGSGVQPMFLTNTFRAINVYVAHVFEESGARPDATKVLV 258
QY 239 LUTDGEKF--GDPLGVEDVIPELDREGVIRVIGVGDAPFRSEKSRQELNTVASKPPDRHV 296
Db 259 IITDGEASDKGNISAAHD-----ITRYIIGIKHFVSVQKTLHI FASEPVEFV 309
QY 297 FOINNFEALTKIQNLREKIFAISGTOTGSSSSFEHEMSQSGPSAAITNSGPLLSTVG 356
Db 310 KILDTFEKLDLFLTELQKIVIEGTSKQDTSFNMLSSGSIADLSRGHVVAVGAVGAK 369
QY 357 DWAGGVF--LYTSKEKSTFINNTRVDSMDNDYLGAAA--IILNRVQSLVLAGAPRYQH 414
Db 370 DWAGFLDLREDLQATFVQBPFLTSQVGGYLGVTYVAMTSSRSRPLLAAGAPRYQHV 429
QY 415 LVAMFR--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTR 472
Db 430 QVLLFOAPEAGRWNTQKIEGTQIGSYFGELCSVDLDODGEAELLIGAPLFGEQRG 489
QY 473 GQVSVCLPRQGRARWCDAYLYGEOGQPMGRFGAALTVLGVDVNGDKLTDVAIGAPGED 532
Db 490 GRVFTY---QRRQSLFEMVSELQDGPYPLGRFGEAITALTIDNGDLTDVAIGAPLEE- 545
QY 533 NRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQ 592
Db 546 -QGVYIIFNGKPG--GLSPQSPQRIQQAQVFGIRWFGRSIHGVKOLGGDLADVVVGABG 603
QY 593 HVLLLRSPVLRVKAIMEFNPREVARNVFECDQVQVVKGEAG-EVRVCLHVQKSTRDLR 651
Db 604 RVVLSSRPVVDVVTLEFSPEEIPVHEVECSYAREEQKGVKUKACPRIKPLTPQ--F 661
QY 652 EQGIQSVVYTDIALDSGRPHSRVNETKNSRRQTVGLTQTCETLKLQLPNCIEDPV 711
Db 662 QGRLLANLSYTLQLDCHMRMRGLFPDGSHELSGNTSITP-DKSCLDLHFPFPIQIDLI 720
QY 712 SPVILRLNPSLV--GTPLSAFGN--LRPVLAEDAQRFLTALPFERNCGNDNICQDDLSI 767
Db 721 SPINVSINFLSEEGTTPROKGRAMQPILRPSIHTV-TKEIPFEKNGCKKKCANLTL 779

A:Accession: A45226
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1151 <BRI>
A:Experimental source: hepatoblastoma cell line HepG2
A:Note: sequence extracted from NCBI backbone (NCBI:P124326)
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F:142-317/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 18.6%; Score 1093.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 8.4e-68;
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

QY 1 FNLDTENAMTFQENARG-FGQSVVLQ---QGSRVVVVGAPQEIIVAAQNRGLSYQCDYSTGS 56
DB 1 FNVVDKNSMTFSGPVEDMEGYTVQQYENEEGKWLLIGSLVGPQKNRTGDVYKCPVGRGE 60

QY 57 CEP-IRLQVPEA-----VNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKG 106
DB 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGQFLACGPLYAYRCGHLHYTTGI 119

QY 107 CFLFGSNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWSTVMEQLKK 166
DB 120 CSDVPTFGVNSIAP--VQECSTQ-LDIVLDGSGNSIYPWDS-----VTAFLNDLLK 170

QY 167 -----SKTLFSLMOYSEEFRIHFTPKFQNNPNRSLIKPITQLLGR-THTATGLRKVV 219
DB 171 RMDIGPKQTVGIVQYGENVTHEFNLNKYSSTEELVAAKKIVQRGRTMTALGTDIAR 230

QY 220 RELFNITNGAKRNAPKILFLTDGKEKFGDPLGVEDVIPELDREGVIRYVIGVDAPR--- 276
DB 231 KEAFTEARGARRGKVKVMVIVTDGESH-DNHRLLKKVIQDCEDENIQRFSAIILGSYNRGN 289

QY 277 --SEKSRQRLNTVASKPRDRHQFNQINNFEALTIQNLREKIPATEGTQSGSSSPEHEM 334
DB 290 LSTEFVBEIKSIASEPTEKHFNFVSDLEALVTIVKTLGERIFALEATADQAAAFEMEM 349

QY 335 SQEGFSAITSNGLLSTVGSVDWAGGVPLYTSKE-----KSTF-INMTRVDSMDNDAYL 388
DB 350 SQTGFSAHYSQDWMLGAVGDMGTVMQXASQIIIPRNTTFNVESKNEPL-ASVL 408

QY 389 GYAAAILNRVQSL-VLGAPRYQHIGLVAMPQNTGMWESANVKGTOIGAYFASLCS 447
DB 409 GYTVNSATASSGDVLYIAGQPRYNTGQVIYRMEDGNKILQTLSGEIGSYFSGILTT 468

QY 448 VVDVDSNGSLDLVLIGAPHY-----YEQTR-GGOVSVCPPLQRCQARWQCD 491
DB 469 TDIKDSDNTDILLVGAPMTYMGTEKEBQGVYVVALNQTRFEYQMSLEPIKQCCSSRQHN 528

QY 492 AVLYGEOQOPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISP 550
DB 529 SCTTENKNEPCGARFGTAAVAKDLNLDGFNDIVIGAPLEDHGGNAVYIYHG-SGKTIK 587

QY 551 SHSQRIAGSKLSPRLQYFGOSLGGODLTWMDGLVDLTVAQGHVLLRSPVLRVKAINME 610
DB 588 EYAQRIPSGDGLTKLFFGQSIHGMEDNGDGLTDVTIGGLGAALFWRSDVAVVAVKVMN 647

QY 611 FNPREVARNVFECDNDVVGKEAG--EVAVCLHVQ-KSTRDLRLEQIQSVVYDIALDS 667
DB 648 FEFKNVIOKKNCH---MEGKETVCINATVCFEVLKSKEDTIYEADLQ-----YRTLDS 700

QY 668 GRPHSRVFNET-----KNSTRROTQVLGLTQTCETLKLQLPNCIEDDPVSPVLRLNFS 721
DB 701 LQIISRSFSGTQERKQVQNRITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDPN 752

QY 722 LVGTPLSAGFNLRPVLAEDAQRLFTALFPFEKNCNGNDNICQDLSITTFMSLDCLVWGG 781
DB 753 LT-DPENG-----PVLDSDSPNSVHEYIPFAKDCGNKKEKICISDLSLHVATTEKDLIVRS 806

QY 782 PRE-FNVTVTRNDGEDSRITQVTFPPFLDLSYRKVSTIQNQRSQBSWLACESASTVE 840
DB 807 QNDKFNVSITVKNTKDSAYNTRTIVHYSPLNLFSGIEATQKO-----SCSEN----- 853

RESULT 8

I45914

Integrin alpha 2 subunit - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999

C:Accession: I45914

R:Kamata, T.; Puzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A:Title: Identification of putative ligand binding sites within the I-domain of integr

A:Reference number: A54402; MUID: 94193647; PMID: 7511592

A:Accession: I45914

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1170 <KAM>

A:Cross-references: GB:L25886; NID:g439695; PIDN:AAB59255.1; PID:g439696

C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F:161-336/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match

Best Local Similarity 18.2%; Score 1071; DB 2; Length 1170;

Matches 333; Conservative 219; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFQ-ENARGFGQSVVLQ---QGSRVVVVGAPQEIIVAAQNRGLSYQCD--DYST 54
DB 19 YNVGLPKAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSFPWSPGPKNRMGDVYKCPVDLST 78

QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKG 107
DB 79 TTCEKLNLTQSTSMNVTMTNMSLGLTLTRNVTGGLTCGLPWAQCGSOYTTTGV 138

QY 108 FLFGSNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEWSTVMEQLK-- 165
DB 139 SDVSPDF-QLRTSFAPAVQTCP-SFIDVVVWCDSENSIYPWD--AVKNFLEXFVQGLDIG 194

QY 166 KSKTLFSLMOYSEEFRIHFTPKFQNNPNRSLIKPITQLL---GRTHATGLRKVRE 221
DB 195 PTKQTQGLIQAANNPRVFNLTNTPKSKD---EMIKATSTQTFQYGGDLTNTFKAIQYARDT 251

QY 222 LFNITNGAKRNAPKILFLTDGKEKFGDPLGVEDVIPELDREGVIRYVIGV-----GDAFR 276
DB 252 AYSTAAGRPAGTKVMVYVVTGESH-DGSKLKAVIDQCNKNILRFGIYAVLYLNRNALD 310

QY 277 SEKSRQRLNTVASKPRDRHQFNQINNFEALTIQNLREKIPATEGTQSGSSSPEHEMSQ 336
DB 311 TKNLKEIKATIASIPIERHFNVSDEADLLEKAGTIGEOIFSIETGVQ-GDNFQNEMSQ 369

QY 337 EGFSAAIT--SNGPLLSTVGSYDWAGGVPLYTSKEKSTFINMT--RVSDMN-DAYLGYA 391
DB 337 EGFSAAIT--SNGPLLSTVGSYDWAGGVPLYTSKEKSTFINMT--RVSDMN-DAYLGYA 391

Db 370 VGFSAEYSPQNNILMLGAVGAYDWSGTGTVQKTPHGLIFSKQAFQEILODRNHSYLGYS 429
Qy 392 AAILNRVQSLVGLAPRYOHIGLVAMFRONTGWESNANV-----KGTQICAYGASL 445
Db 430 VASISTGNSVHFVAGAPRANTYQGVLYSYN-----ENGNTVVIQSGDQIGSYGSLV 484
Qy 446 CSVDVDSNGSTDLVLGAPHYEOTR--GGQSVCPPLRGQARWQCDVAVLYGECQCPWG 503
Db 485 CAVDVNKDTITDVLVAGPMYNDLKKEGRVYLFITKG-ILNWH--QFLEPGNLENA 541
Qy 504 RFQAAITVLGDVNGDKLTDVAIGAGEEDNRGAVYLFHGTSGSGISPSHSQRIAGS--KL 561
Db 542 RFGSAIAALSDINMGDNFVIGSGPLENQNSGAVIYNGHEGM-IRLRYSQKILGSDRAF 600
Qy 562 SPRLQYFGQSLGGQDLMGLVDLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVF 621
Db 601 SSLHVFGRSLDGYDGLNGDSITDVSAGFQVQVQLWSQSIADVSVDASTPKKI--TL 658
Qy 622 ECNDQVVGKEAGEVRVCLHVQKSTRDLREGQIQSVVYDLDL-----SGRPHSAFVN 677
Db 659 NKAAEI-----KLKCF-----SAKFRPTNQNNQVAIVYNIIDEDQFSRVISGLFK 707
Qy 678 ETKNSTRQTOVLGLTQTCB--TLKQLPNCIEDPVSPVILRLNFSL--VGTPLSAFNL 733
Db 708 ENNERCLQKTMIVSQAQRCSEYIIHQEPS---DIISPLNLCWNISLENPGT----- 756
Qy 734 RPLVAEDAQRLFTALPFKCNKCNNDICQDLDLITP-----SPMSLCLVGGPRENTV 789
Db 757 NPALAEAYSETVKVFSIPFHKDCGDGVCISDLVNLVQQLPATQOQPFIVSNQNKRTFSV 816
Qy 790 TVRNDGDSYRTOVTFPPFLDLSYRKVSTLQNRQSORQSRWELACESAST--EVSGLKSTS 848
Db 817 OLKKNKESAYNTEIVDFSENLF-----ASWMPVDGTEVTQCIASSOKSVT 864
Qy 849 CSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTETFOLELPVKY 908
Db 865 CNVGPALKSKQQTFTINFDFNLQ-NLQOASISPRALSESQENMADNSVNLKLSLY 923
Qy 909 AVYMWVTSHGVSUKYLNFTASNTSRVMOHQYQVSNLQOR-----SLPISLVFLV 958
Db 924 DAEIHIT-RSTNINFYEVSLDGNVSVV-HSFE--DIGPKFIFSKVTGTSVSPVMA--- 976
Qy 959 PVLNQTIVMDRPQVTFSEN---LSSSTCHTKE-----RLPSHSDFLAE- 998
Db 977 -----SVIITHIFOYTKDKPLMYLTGVHTDQAGDISCEAEINPLKIGQTSVSPKSEN 1030
Qy 999 LRKAPVNCISAVCQRIQCDIPFPGIOEEFNATLKNLSFDWYIKTSHNHLIVSTAEI- 1057
Db 1031 FRHIKELNCRTASCNIMCMLRDLQVKGVEFLNVSTRIWNGTFAASTFQTVQLTAAAEID 1090
Qy 1058 LFNDSVFTL-----LPGQAFVRSQTETKVEPPE-VPNPLPLIVGSSVCGLLILLALITA 1110
Db 1091 TYNPQIYVIBENTVIP-----LTIMKPEKVEVPTGVIVGSIAGILLALLA 1140
Qy 1111 ALVKLGFFKRYQKDM 1125
Db 1141 ILWKLGFFKRYKXDM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze

submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A:Cross-references: EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473099
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F:169-344/Domain: von Willebrand factor type A repeat homology <WVA2>

Query Match 18.0%; Score 1057; DB 2; Length 1178;

Best Local Similarity 27.8%; Pred. No. 3.1e-65;

Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;

Qy 1 FNLDTENAMTFQ-ENARGFGOSVVL-----QGSRVVVGAPQEIIVANQSGSLYQC---DYST 54
Db 27 YNVGLPGAKIFSGPSSQEGYSVQQLTNPQGNWLLVGSPPFENRMGDVTKCPVDLPT 86
Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPQQLACGPTVHOTCSSENTYVKGIC 107
Db 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTNPCTGGFLTCLPLWAHQCGNQYATGIC 146
Qy 108 FLFGNLRQOPQ--KFPFALRGCPQEDSDIAFLIDGSGSIIPHFRMKWVSTVMEOL 164
Db 147 ----SDVSPDFQFLTFSFPAVQACPSL-VDVVVVWCDSENSIYP--WEAVKNFLVKFVTGL 199
Qy 165 K--KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPITQLLG-RTHATATGLRKVVRE 221
Db 200 DIGPKTKQVALIQYANEPRIIIFNLNDFETKEDMVOATSETROHGGDLTNTFRAIEFADY 259
Qy 222 LFNITNGARKNAFKILFLITDGEKFGDPLGYEDVIPELDREGVIRYVIGV-----GDAFR 276
Db 260 AYSQTSGGPAGATKVMVVYTDGESH-DGSKLTVIQCCNDDEILREGIAVLGLNRLNLD 318
Qy 277 SEKSRQELMTVASKPRDRHVFOINNFEALKTIONLREKIFAIEGTQTSSTSSFEHMSQ 336
Db 319 TKNLIKEIKAIATASTPTERYFFNVADAEALLKAGTLGEQIFSIETGVQG-GDNFQEMEAQ 377
Qy 337 EGESA--AITSNGPILLSTVGSVDWAGCVPLYTSKESKSTFINMT--RVDSDMN-DAYLGVA 391
Db 378 VGFSADYAPQNDILMLGAYGAFDWSGLTQVETSHKPVIPFQKAFQDVLQDRNHSFLGYS 437
Qy 392 AAILNRVQSLVGLAPRYOHIGLVAMFRONTGWESNANV---KGTQICAYGASLCS 447
Db 438 VAAISTEDGVHVFAGAPRANTYQGVLYSYN---QGNVTVIQSHRGDQIGSYFGSVLCS 494
Qy 448 VVDVDSNGSTDLVLIGAPHYYEOTR--GGQSVCPPLRGQARWQCDVAVLYGECQCPWG 505
Db 495 VDVQDQITIDVLVGAPTYMNDLKKEGKVYLFITKGLNQHQ---FLEGEPTGNARF 551
Qy 506 GAALTIVLGDVNGDKLTDVAIGAGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPR- 564
Db 552 GSAIAALSINMGDNFNDVIVGSEVENENSGAVIYNGHQGT-IRTKYSQKILGNSGAFRR 610
Qy 565 -LQYFGQSLGGQDLMGLVDLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFEC 623
Db 611 HLQFFORSIDGYDGLNGDSITDVSIGALGOVQLWSQSIADVAIEALFTP----- 660
Qy 624 NDQVWKGKAGEVRVCLHVQKSTRDLREGQIQSVVYDLDL-----SGRPHSAFVN 679
Db 661 -DKITLLNKDAKITLKLCPRAEFPRPAGQNNQV--AILFNMTLDADGHSRVRTSRGFREN 717
Qy 680 KNSTRQTOVLGLTQTCET--LKLQPLNCIEDPVSPVILRLNFSLVTGTPLSAFGNLRPVL 737
Db 718 SERFLQNMVVADEVQKSEHHSIQKPS---DVNPLDLRVDISLENPGTS-----PAL 768
Qy 738 AEDAQRLLTALPFPEKNCNDNICODDLISI-----TFSPMSLCLVVGPPREFNVT 788
Db 769 EAYSETVKVFSIPFYEKCGSDGICISDLIDVQQLPAIQTSF-----IVSNQNKRLTFS 823
Qy 789 VTVRNDGDSYRTOVTFPPFLDLSYRKVSTLQNRQSORQSRWELACESAST--EVSGLKST 847
Db 824 VILKNRGESAYNTVLAEFSENLF-----ASFMPVDQTEVTCEVSGSQSV 871
Qy 848 SCINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTET--FOLELP 905
Db 872 TCDVGYPALKEQOQVTFITDFNLQ-NLQOAAINFPQAFSESQ--ETNKADNSVSLTIP 928
Qy 906 VKYAVVMVTSHGVSUKYLNFTASNTSRVMOHQYQVSNLQOR-----SLPISLV 955


```

Db 929 LLYDAELHLT-RSTNINFEISSDENAPSVIK--SVEDIGPKFFISLKVTAGSAPVSWA 984
QY 956 FLV-----PVRNQTIVMDRPPQVTP-SENLS 980
Db 985 LVTHIHPQVTEKKNPLLYLTGTDQAGDISCTABINPLKLPHTA-----PSVSFKENFR 1040
QY 981 STCHTKRLPSHSDFLAELRKAPVNCSTAVCORIQCDIPFFGIEEENATLKGNSLSDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMAEYFINVTRVWMT 1080
QY 1041 YIKTSHNHLIVSTAEILFNDSVFTLLPQCGAFVRSOTETKVEPEVNPPLIVGVSVG 1100
Db 1081 FAASTFQTVLTAAAEIDTHNPOLFEIBENAVTIFLMIMKPEKAEVPT--GVIGSIIA 1138
QY 1101 GLLILLALITAAALYKLGFKRQYKDM: 1125
Db 1139 GILLLLAMTAGLWKLGFGRQYKDM 1163

RESULT 10
A35854
Integrin alpha-1 chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 20-Sep-1999
C:Accession: A35854; S11243
R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Bach, F.; Carbonetto,
J. Cell Biol. 111, 709-720, 1990
A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin A
A:Reference number: A35854; MUID:90338125; PMID:2380249
A:Accession: A35854
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1180 <IGN>
C:Cross-references: GB:X52140; NID:956493; PIDN:CAA36384.1; PID:956494
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C:Keywords: Cell adhesion; cytoskeleton; transmembrane protein
F:170-345/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 17.9%; Score 1055; DB 2; Length 1180;
Best Local Similarity 27.4%; Pred. No. 4.3e-65;
Matches 346; Conservative 196; Mismatches 475; Indels 246; Gaps 48;

QY 1 ENLDTENAMTFOENARG-FGQSVWL---QGSRVVUGAPQEIIVAAVQNRGSLYQCDYSTGS 56
Db 29 FNVVDKNSMFSFGPVEDMFGYTVQOYENEEGKMWLIGSLVPGPKARTGDVTVKCPVGRER 88
QY 57 CEP-IRLQVPVEA-----VNMSLGLSLAATTSPOLLACGPTVHQTCSENTYVXGL 106
Db 89 AMPCVKLDLPVNTSIPNVEIKENMTFGSTL-VTPNPGGFLACGPLYAYRCOHLHYTGI 147
QY 107 CFLFGSNLRQOPKPEALRGCPQBDSDIAFLIDGSGSIIPHDFRMKEWSTVM---EQ 163
Db 148 CSDVSPTFOVNSFAP--VQECSTQ-LDIVILDGNSIYP-----WESVIAFLNDL 196
QY 164 LKK-----SKTLFSLMYSSEERIHFTKEFQNNPNRSLIKPITQLLG-RHTHTATLGRK 217
Db 197 LKRMIDGPKQTQVQVGENYTHBFNLNKSSTEEVLVAANKIGRQGLQMTALGIDT 256
QY 218 VVRELFTNNGARKNAFKILFLTDGKFGDPLGVEDVPELDREGVIRYVGVGDAPFR- 276
Db 257 ARKEAFTARGARGVKVMVITDGESH-DNYRLKQVTDCEDENIQRFSTAILGHVNR 315
QY 277 ----SEKSRQELNTVASKPPRDHVFQINNFEALKTIQNLRKIFAIEGTQTGSSSEFH 332
Db 316 GNLSTEKFEVEIKSIASEPTEKHFFNVSDALVITVKAIGERIFALEATADQSAASFEM 375
QY 333 EMSQGFSAIYNSGFLSTVCSYDWAGVFLYTSKESKTFINMT--RVDSDMD--AY 387
Db 376 EMSQTGFSAHQDMMVLGAVGAYDWMGTVMQKAMQVPIPHNTTFQTEPAKKNPLASY 435
QY 388 LGYAAAAILLRNVQSLVGLGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCS 447

```

```

Db 436 LGYTVNSATIPGDVLYIAGQPRVNHHTGQVVIYKMGEDGNINILQTLGGEQIGSYFGSVLTT 495
QY 448 VDVDSNGSTDLVLIGAPHY-----YEOTR-CGQVSVCPPLRQGRARWQCD 491
Db 496 IDIDKDSYTDLLLVGAPMYMGTEKEBOGKVYVYVAVNQTRFYQMSLEPIRQTCSSSLKDN 555
QY 492 AVLXGEOQPMWG-RFGAALTVLGVNDGDKLTDVAIGAPGEEDNRGAVLVFHTSGSGISP 550
Db 556 SCTKENKNEPCGARFGTAIAAVKDLNVDFNDVIGAPLEDHAGAVVIYHG-SGKTIRE 614
QY 551 SHSRIAGSKLSPRLQVFGQSLTMDGLVOLTVGAGOHVLLLRASQPVLRVKAIME 610
Db 615 AYAQRIPSGDGDKLTKFQGGIINGMDLNGDGLDVTGIGLGAALFWARDVAVVVKVTMN 674
QY 611 FNPBEVARNVPECNDQVVKGEAG-EVRVCLHVQ-KSTRDLRREGQIQSVVTVVDLALDS 667
Db 675 FEPKNVNIQKNCR---VEGKETVCINATMCFHVKLSKESDIYEADLQ---YRVTLDS 727
QY 668 GRPHSRVAFNET-----KNSTRQTQVLGLTQTCETLKLQLPNCI-----EDPVS 712
Db 728 LRQISRFFSGTQERKIQRNITVRESE-----CIRHSPYMLDKHDFQD 770
QY 713 PIVLRNPSLVGTPLSAPGNLRPVLAEADAQLFTALPFPFKCNCGNDNICQDLSITTFSM 772
Db 771 SVRVTLDFNLTDPENG-----PVLDDALPNSVHEHIPFAKDCGKERCISDLTLNVSTT 824
QY 773 SLDCLVGGPRE-FNVTVTVRNDGEDSYRTQVTFPFFPLDLSYRKVSTLQNRQSQRWLA 831
Db 825 EKSLLIVKSHQHDKFNLSLTVKNGKDSYNTTRTVQHSNPLIFSGIEEIQKD-----S 876
QY 832 CESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLL-LKANVTSEN 890
Db 877 CESN-----QNITCRVGYPLRAGETVTFKLIIFQFNTSHLSENAIHLSATSDSEE 927
QY 891 NMPTNKTETFOLELPVKYAV---YMWVTSHGVT-----KYLNTFASNTSRVMQHQ 939
Db 928 PLESINDNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINV 985
QY 940 YQVSNLQORSIP---ISLVF-----LVPVRLNQTIVWD-----RP----- 971
Db 986 YTKRRGHFPMPELQLSISFPNLFTADGVPVLYPIG-----WSSSDNVNCRPSLEDVFG 1039
QY 972 ----QVTFS-----ENLSSTCHTKERLPSHSDFLAELRKAPVNCVCSIAVCRIQ 1016
Db 1040 INSGKMTISKEVLKRGTIQDCSTC-----GVATITCSLLPSDLSQ 1082
QY 1017 CDI-----PFFGIOEEF---NATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLL 1067
Db 1083 VNVSLLLWKPTF-IRAHFSSNLTLRGELK-----SENSLTLSSN----- 1123
QY 1068 PQGAFVRSQETKVEPEVNPPLP---IVGSSVGGILLLLALITAAALYKLGFFKRYKDM 1125
Db 1124 -----RKRELAIQISKDGLGPRVPLWILLISAFAGLLLLMLLILALWKLKIGFFKRLKKK 1177
QY 1126 MSE 1128
Db 1178 MEK 1180

```

RESULT 11

A33998

Integrin alpha-2 chain precursor - human

N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2

C:Species: Homo sapiens (man)

C>Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 21-Jul-2000

C:Accession: A33998; B56793; A53117

R:Takada, Y.; Hemler, M.E.

J. Cell Biol. 109, 397-407, 1989

A:Title: The primary structure of the VLA-2/collagen receptor alpha (2) subunit (platelet)

A:Reference number: A33998; MUID:89308879; PMID:2545729

A:Accession: A33998

A:Molecule type: mRNA

A:Residues: 1-1181 <TAK>

A;Cross-references: GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Catalmel, B.; Parmentier, S.; Leung, L.B.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GP1a, GP1c, GP1c*, GP1a and GP1b) from platelets
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 18-21 <TUT>
A;Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;1105, 112, 343, 432, 460, 475, 699, 1057, 1074, 1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 17.9%; Score 1054; DB 2; Length 1181;
Best Local Similarity 26.7%; Pred. No. 5.1e-65;
Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;

Qy 1 FNLDENAMTQ-ENARGFGOSVVQL---QGSVVVVGAPQEIIVAAQNGSLYQC--DYST 54
Db 30 YNVLGPEAKIFSGSSSQFGVAVQFIPNKGWLLVSGSPGFENRMDGVKCPDLST 89
Qy 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATSPQALLACGPTVHTQTSNTYVKGLC 107
Db 90 ATCEKLNQTSIPNVTETKNTNLSGLILTRNMGTLGFLTCGPLWAQCCQNGQYTTGVC 149
Qy 108 FLFGSNLRQOQKPEALRGCPQSDSDIAFLIDSGSIIIPHDFRMKEWSTVMEQLK-- 165
Db 150 SDISPDF-QLSASFSPATQPCPSL-IDVVVVCDESNSIYPWD--AVKNFLEKFFVQGLDIG 205
Qy 166 KSKTLFSLMOYSEEFRIHFTKFEQNNPNRSLIKPITQLLG-RTHATGLRKYVRELFN 224
Db 206 PTKTQVGLIQANNPRPVVFNLTNYKTEEMIVATQTSQYGGDLTNTFGALQYARKAYS 265
Qy 225 ITNGARKNAKILFLTLTDEKFGDPLGYEDVIPELDRGVIRYVIGV-----GDAPRSEK 279
Db 266 AASGRRSATKVMVVVTDGESH-DGSLMKAVIDQCNHDLNIRFGIAGVLGLNRLALDTKN 324
Qy 280 SRQELNVTASPPRHHVQIINNFEALKTIQNLREKIFAIGTQTGSSSSSEHEMSQEGF 339
Db 325 LIKEIKATASIPTRYFFNVSDAALKEKAGTLGEQIFSIETVQG-GDNFQMSQVGF 383
Qy 340 SAATISNGP--LLSTVGSYDAGGVFLVTSKEKSTFINMT--RVDSDMN-DAYLGAAAI 394
Db 384 SADYSQNDILMLGAVGAFSGTGVTKTSHGLHLPKQAFDQILQDRNHSYLGYSVAA 443
Qy 395 ILNRNVQSLVGLGAPRYQHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLCSV 448
Db 444 ISTGESTHFVACPRANTVQGVLYSVN-----ENGNITVIAHRGDDQIGSYFGSLCSV 498
Qy 449 DVDSNGSTDLVLIGAPHYEQTR--GGQSVVCLPRGQARWQCDVLYGGQGPWGRFG 506
Db 499 DVDKDTITDVLVAGPMVMSDLKEEGRVYLPITKKILGQHQ---FLEGPEGIENTRFG 555
Qy 507 AALTVLGDVNGDKLTDVAIGAPGEEDNRGAVFLPHGTSGSGISPSHSQRIAGS--KLSPR 564

Db 556 SAIAALSNDIMDGFNDVIVGSPLENQNSGAVIYNHQOT-IRTKYSKILGSDGAPRSH 614
Qy 565 LOYFQSGISGGODLTWGLVDLTVCAGQGVLLLRQSVLRVKAIMEEPREVARNVFECN 624
Db 615 LOYFGSLDGYGDLNGDSITDVISGAGVOVQLWSQSTADVAIEASFTEPKI--TLVKN 672
Qy 625 DQVWKGKEAGEVRVCLHVQKSTRDLRREGQIQSVVYTYDLALD-----SGRPHSRVAFNETK 680
Db 673 AQII-----LKLCF---SAKPRTKQNNQVAIYNNITLDADGFSRVRTSRGLFKENN 721
Qy 681 NSTRQQTQVLGUTQTC--ETLKLQPNCIEDPVSFIVLRNLSLVTGTPLSAAGNLRPVLA 738
Db 722 ERLCKQNVVNAQSCPEHIIYIQEPS---DVNSLDLRVDSLENPGTS-----PALE 772
Qy 739 EDAQRLFTALFPFEKNCNDNICQDDLSTIF---SFMSLCLVVGPPREFNVTVTRND 794
Db 773 AYSETAKVSIPIHKDCGEGDGLICISDLVDVTRQIPAAQEQPIFVSNQKRLTFSVTLKNK 832
Qy 795 GEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVVDFSENLPF-----ASFSLPVDGTEVTCQVAASQKSVACDVGY 880
Qy 854 PIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTQFQLELPVKVAVMV 913
Db 881 PALKRQQTTFINFDNLQ-NLQNASLSFOALSESQEEKADNLVNLKIPULLYDAEI- 938
Qy 914 VTSHGVS TKYLNFTASENTSRVMQHOYQVSNLQOR-----SLPISLVFLV---- 958
Db 939 ---HLTRSTNINPIEISSDGNVPSIVHSFEDVGPKEFISLKVTTGSPVPMATVIIHQ 995
Qy 959 -----PVRINQTVIWDPRQVTF-SENLSSTCHTKER 988
Db 996 YTKENKPLMYLTCVQTDKAGDISCNADINPLKIGQT---SSSVSPKSENFR---HTKE- 1047
Qy 989 LPSHSDFLAELRKAPVNVCSIAVCQRIQCDIPFGIQEFNATLKNLSFDWYIKTSHNH 1048
Db 1048 -----LNCRTASCNVTCMLKDVHMKGBYFVNVVTVTRINWGTFFASSTFQT 1091
Qy 1049 LLIVSTAET-LFNDVSTETLLPGQAFVRSQTEKVEPFPNP-----LP--LIVGSSVG 1100
Db 1092 VQUTAAEINTYNEPIYVI-----EDNTVTIPLIMKPDKEAEPVTVGIISIIA 1141
Qy 1101 GLLLALITAAALYKLGFFRQYKDM 1125
Db 1142 GILLALLVALKLGFFRQYKDM 1166

RESULT 12
A41131
Lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N;Alternate names: integrin alpha-4
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999
C;Accession: A41131; S16742
J;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzman, B.; Weisman, I.L.
J. Cell Biol. 115, 1149-1158, 1991
A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
A;Reference number: A41131; MUID:92064645; PMID:1840602
A;Accession: A41131
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1039 <NEU>
A;Cross-references: EMBL:X53176; NID:g51484; PIDN:CAA37316.1; PID:g51485
C;Superfamily: integrin alpha-4 chain
C;Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 663; DB 2; Length 1039;
Best Local Similarity 22.7%; Pred. No. 9.3e-38;
Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

Qy 1 FNLDENAMTQ-ENARGFGOSVV-QLQGR--VVGAPQEIIVAN-----QRGLYQCDY 52
Db 1 FNLDENAMTQ-ENARGFGOSVV-QLQGR--VVGAPQEIIVAN-----QRGLYQCDY 52

Db 41 YNLDPENALLYQGPSGTLFGYSVVLHSHGSKRWLIIVGAPTASWLSNVAIVRCGI 100
QY 53 STG---SCPIRLQVP-----VEAVNMSLGLSLAATTSP-PQLLACG---PTVHQ 95
Db 101 RKNPNOTCQLQSGSGSGPCGKTCLEBRDNQWLVTLRQPGENGSIYTCGRWKNYFY 160
QY 96 TCSNTYVYKGLCLFGLFSGNLRQQPKPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKE 155
Db 161 MKSDNKLPTGICYVMPSDLRTLSK-----185
QY 156 WSTVMEQLKSKTSLFSLMQYSEPRIFHTFKFQNNPNRSLIKPITQLLGRTHATGL 215
Db 186 -----RMAPCYKDYT-----195
QY 216 RKVVRELFNITGARKNAFKILLTDGEGKFGDPLGVEDVIPELDREGVIRYVIGVDAP 275
Db 196 -----RKFGS-----200
QY 276 RSEKSRQELNTVASKPPRHVFOINNFEALKTIQNLREKIFAIEGTQTGSSSFFHEMS 335
Db 201 -----NFAS-----C 205
QY 336 QEGFSAITNSGPLLSTVSGSYDWAGGVFLY--TSKEKSTFINMTRVDSMDNAYLGYA-- 391
Db 206 QAGISSFYTDLIWVGAPSSYWTGTVFVYNIITNQYKAFVD--RQNVKFGSYLGSYVG 263
QY 392 AAILRNVSQSLVLGAPRYGHICGLVAMFRQNTGMESNANV-----KGTOLGAVFGSLCS 447
Db 264 AGHFRSPHTTEVVGAPQHQIYKAYIF-----SIDENELNIVEMKGLGSLYFGASVCA 319
QY 448 VDVSNGSTDLVLIGAPHYEQTRGGQVSVCPLPRGQRA--RMQCDVAVLXGEOQPMGRFG 506
Db 320 VDLNADGFSDL-LVGAPMQSTIREGRVFYV-INSGMGAVVMEVRLVSGDKYA-ARFG 376
QY 507 AALTVLGVNNGDKLTVAIGAPEDNRGNAYVLFHTSGSGISPSHSQRIAGSKLSPRIQ 566
Db 377 ESTANLGIDINDGFDIAGPQEDDLRGAVIYNGRV-DGISSTYSQRIEQGQISKSLR 435
QY 567 YFGQSLSGGODLMDGLVGLTVCA--QGHVLLRSOPVLVRVKAIMEFNEVARNVFEEN 624
Db 436 MFGQSLSGQIDADNGVDVAVAFQSDSAVLLRTPVIVVEASLS-HPESVNRTRKFDCT 494
QY 625 DVVVKKEAGEVRVCLHVQKSTRDLREGIQIQVWVYDLALDSGR---PHSRAVF--NET 679
Db 495 -----ENGLPSVCMHLTLCFSYKGVPGYIVLVFNVSLLDVHRKAESPSRFYFSNGT 547
QY 680 KNSTRQTVLGTQTCETLKLQLPNCIEDVPSPVILRNFSL-----VOTPLSAFNGLR 734
Db 548 SDVITGSIRVSSSGEKRTHQAFMRKDVDRDILTPHVEATYHLGHHVITKRNTTEPPPLQ 607
QY 735 PVLAEDAQR-LFTALPFEKNCNGDNICODDLSITFS-----FMSLDCLVVGPPREFNV 787
Db 608 PILQKKEKDVIRKMINARFCAYEN-CSADLQVSARKVGLPKPYENKTYLAVGSMKTIWL 666
QY 788 TVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWSRLACASASTEVSGALKST 847
Db 667 NVSLFNAGDADAYETLLNVQLPTGLYFIKILDEEK-----QINCE---VTSSGIVK-L 716
QY 848 SCISNHPPIPFENSEVTNITFDVDSKASLGNKLLKANVTSEN--NMPRTNKTETFOLELP 905
Db 717 ACSGLGIYVDRLSRIDISFLDDVSSLSRAHEDLSISVHASCENEGELDQVRDRNRTLTP 776
QY 906 VKYAVVYVTSHV--STKYNLTASNTSRVMQHO-----YOVSNLQORSLP-ISLVEL 957
Db 777 LRYEV--MLTVHGLVNPSTGFVYGSSENEPETCMAEKLNLTPHVTINTGISMAPNVSVKIM 834
QY 958 VPVRLNQTIVDRP--QVTFSENLSTCTKE-----RLPSHSDFLAELR 1000
Db 835 VP---NSFLPQDDKLFNVLDOVTTGQCHFKYHGRECTFAQQKGIAGTLTDIVKFLSKTD 891
QY 1001 KAPVAVNCSTIA--VCQRIQCDIPFFGQEEFNATLKGNLSPDWYIKTSHNHLITVSTAEIL 1058
Db 892 KR-LLYCMKADQCHLDFLCN---FGKMESG-----KEASVHIQLEGPSIL 933

QY 1059 FNDVSFTLLPGQGAFAVRSQTETKVFPEFVNP-----1090
Db 934 ENDETSSL-----KFEIKATAFPEPPEPKVIELNKDENVAHVFLGLELHQRKRHF 983
QY 1091 -LPLIVGVSVOGLLLALITAAALYKLGFFKQYKMMSE 1128
Db 984 TIIITITISLLGLVILLISCVMMKAGFFKQYKSILOE 1022
RESULT 13
S06046
N:Altein alpha-4 chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 24-Sep-1999
C:Accession: S06046; A39355; D28018
R:Takada, Y.; Ellices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other int
A:Reference number: S06046; MUID:89356603; PMID:2788572
A:Accession: S06046
A:Molecule type: mRNA
A:Residues: 1-1038 <TAK>
A:Cross-references: GB:X16983; EMBL:X15356; NID:G33945; PIDN:CAA34852.1; PID:G33946
R:Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A:Title: Characterization of the alpha-4 integrin gene promoter.
A:Reference number: A39355; MUID:91239513; PMID:2034655
A:Accession: A39355
A:Molecule type: DNA
A:Residues: 1-93 <ROS>
A:Cross-references: GB:M62841
R:Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A:Title: The very late antigen family of heterodimers is part of a superfamily of mole
A:Reference number: A94151; MUID:87204112; PMID:3033641
A:Accession: D28018
A:Molecule type: protein
A:Residues: 40-50, E' 52-53 <TA2>
C:Genetic: B
A:Gene: GDB:ITGA4; CD49D
A:Cross-references: GDB:128032; OMIM:192975
A:Map position: 2q31-2q32
C:Superfamily: integrin alpha-4 chain
C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; trans
F:1-39/Domain: signal sequence #status predicted <SIG>
F:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>
Query Match 10.9%; Score 642; DB 2; Length 1038;
Best Local Similarity 23.8%; Pred No. 2,8e-16;
Matches 293; Conservative 175; Mismatches 409; Indels 356; Gaps 53;
QY 1 FNLDTENAMTFQ-ENARGFGQSVV-QLOGSR--VVVGAPQEIIVAAAN---QRGSLYQCDY 52
Db 40 YNVDTESALLYQGPHTLFGYSVVLHSHGCAWLLVGAPTANWLANASVINPGAIYRCRI 99
QY 53 STG---SCPIRLQVPVEAVNMSLGLSLAATTSPQQLACGPTVQTCSTYVYKGLCLFL 109
Db 100 GKNPGQTCBQLQSGP-----NGEP---CG---KTCLEERDNOWL---133
QY 110 FGSNLRQOPQKPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEWVSTVMEQLKSKXT 169
Db 134 -GVTLRSQGE-----NGSIVTCGHR---W-----KN 156
QY 170 FLSLMQYSEPRIFHTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGA 229
Db 157 IFVI-----160
QY 230 RKNAPKILFLTDGSKFGDPLGVEDVIPELDREGVIRYVIGVDAPFRSEKSRQELNTVAS 289
Db 161 -KNENK---LPTGCGYGP-----PDL-----RTELSKRIA 187

Qy 290 KPPRDHVFQI--NNFEALKTIQNLREKIFAETGTGTGSSSSFEHMSQEGFSAAITSNP 348
Db 188 PCYQDYVKKGFEAS-----CQAGISSFTYTKDLI 217
Qy 349 LLSTVGSYDAGVFLY--TSKSKSTFINMTRVDSMDNDAYLGYA--AAIILNRVQSL 403
Db 218 VMGAPGSSYMTGSLFVNITTKYKAFDKQNKVFP---GSYLGYSYGAGHFRSQHTTEV 274
Qy 404 VLGAAPRYOHLGLVAMFRQNTGMWESNANV---KGTGIGAVFGASLCSVDVDSNGSDLV 459
Db 275 VGGAPQHEQJGKAVIF---SIDEKELNILEMKGKGLGSYFGASVCAVDLNAUGFSDL- 329
Qy 460 LIGAPHYYEOTRGQGVSCPLPRGORARWQC--DAVLYGEOQGPGRFGAALTIVLGVNGD 518
Db 330 LVGAPMQSTIREGRVFY--INSGGAVMAMETNLVGSOKYA--ARFGESIVNLGDDND 387
Qy 519 KLTVAIGAPGEEDNRGAUYLFHGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDL 578
Db 388 GFEDVAIGAPQEDDLOGAIYIYNGRA--DGISSTFSQRIEGLQISKLSMFQQSISGIDA 446
Qy 579 TMDGLVDLTGCA--QGHVLLRSOPVLVRVKAIMEFNPVARNVFCNDQVVGKEAGEV 636
Db 447 DNNGYVDVAGAFRSDSAVLTRPVIVDASLS--HPESVNRTRKFDG-----VENGW 498
Qy 637 RVCLHVQKSTRDRREGIQSVVTYDLALDSGR-----PHSRVFNETKSTRRTQTVLG 691
Db 499 SVCIDLTLTCSYKKEVPGYIVLFYNSLSDVNRKAESPFRPYFSSNGTSDVITSGIQVSS 558
Qy 692 LTQTCETLKLQLPNCIEDPSPVILRNFSLVGTPLS-----AFGNLRPVLAEDAQR-LF 745
Db 559 REANCRTHQAFMRKDVDRDILTPQIETAAHYHLGPHVISKRSTEEPPPLQILOQKKEKIM 618
Qy 746 TALPFEKNCNDNICQDDLSIT--FSFM-----SLDCLVGPGREFNVTVTVRNDGDSY 799
Db 619 KKTINFARCAHEN--CSADLOVSAKIGFLKPHENKTYLAVGSMKTLMLNLSLFWAGDDAY 677
Qy 800 RTQVTFPPDLRYKRVSTLQNRQSRWSRLACSSASTEVSGALKSTSCSINHPIPPEN 859
Db 678 ETTLHVKLPGVLYKILEBEK-----QINCE---VTDNSGVVQ--LDCSIGIYVDHL 727
Qy 860 SEVTNFTTFVDVSKASLGKLLKANVTSEN--NMPRTNKTFOLELPVKVAVYVMTSH 917
Db 728 SRIDISFLDVSLSRAEDLSITVHATCENEEDNDLKHGRVTVVAIPLKYEVKLVHGF 787
Qy 918 GVSTKYNLFTASENTRV-----MOHQYQVSNLQORSIP--ISLVLVPLVRLNQTVINDRP 971
Db 788 VNPTSFVYGSNDENEPETCMVEKMNLTTPHVINTGNSMAPNVSEIWPNSFS-----P 840
Qy 972 QVTFSENIL-----SSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIPFGIQ 1025
Db 841 QTDKLFNILDVQTITGECFENYQ-----RVCALEQ-----Q 872
Qy 1026 EEFNATLKGNSLSPWYIKTSHNLLIVSTAE-----ILFN-----DSVFTLLPQGG 1071
Db 873 KSAMQTLKGIYRF---LSKTDKRLLYCIKADPHCLNFCNFGKMSGKEASVHIQLEGR 929
Qy 1072 AFVRSQRTKVEPEV-----PNP-----LPLIV 1095
Db 930 S-ILEMDSALKFEIRATGPPEPNRVIELNKDENVAVHLLGHQRPKRYFTTIVIS 988
Qy 1096 GSSVGGLLLLALITAAALYKGLFFKRYQKDMSE 1128
Db 989 SLLGLGLVILLISVVMKAGFFKRYKSILOE 1021

RESULT 14

158409
Integrin alpha-9 chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2000
C:Accession: I58409; A49459
R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994

A:Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in sm.
A:Reference number: I58409; MUID:94119603; PMID:8290272
A:Accession: I58409
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1035 <RES>
A:Cross-references: GB:D25303; NID:G464180; PIDN:BAA04984.1; PID:G533327
R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A:Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel par
A:Reference number: A49459; MUID:94064789; PMID:8245132
A:Accession: A49459
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 30-1035 <PAL>
A:Cross-references: GB:L24158
C:Superfamily: integrin alpha-4 chain
C:Keywords: glycoprotein; metal binding; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7% Score 630; DB 2; Length 1035;
Best Local Similarity 26.5% Pred. No. 1.9e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;

Qy 331 EHMSSQEGFSAAITSNGLSTVGSYDAGVFLYTSKEKSTFINMT-RVDSMDNDAYLG 389
Db 190 EHSGCQAGIAGFFTEELVVMGAPGSGFYWAGTIKVLNLT-D-NLYLKLNDVIMNRRYTYLG 248
Qy 390 YA--AAIILNRVQSLVGLGAPRYQHIGLVAMFR--QNTGMWESNANVKTQIGAYFASL 445
Db 249 YAVTAGHFSPSTIDVVGGAPODKIGKVIYFADRNRSGTLIKIFQASGKMGSGYFGSSL 308
Qy 446 CSVDVDSNGSTDLVLGAPHYYEOTRGQGVSCPLPRGORARWQCDAVLYGEOQGPWGRF 505
Db 309 CAVDLNGDGLSDL-LVGAPMFSIRDEGQVTVY-INRNGALAE-QIALTGD-CAYNAHF 364
Qy 506 GAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAUYLFHGTSGSGISPSHSQRIAGSKLSPRL 565
Db 365 GESIASLDDLNDGDFDVAIGAPKEDDFAGAVYIYHGDAG-GIVPQYSMKLSGQKINPVL 423
Qy 566 QYFGQSLSGQDLTMDGLVDLTGCA--QGHVLLRSOPVLVRVKAIMEFNPVARNVPEC 623
Db 424 RMFGQSISGIDMDGNGYPODVTVGAFMSDSVLLRARPVITVD-VSIFLPGSINITAPQC 482
Qy 624 NDQVVGKEAGEVRV-CLHYQ-----KSTRDLREGQIQSVVTVYDLALDSGRPHGRAV 675
Db 483 HD-----GQPPVNCINVTTCESFHGKHVPEI---GLNYVLMAVAKKEKQMPRVY 531
Qy 676 F---NETKNSTRRTQTVLGTQTCETLKLQLPNCIEDPSPVILRLNFSL-----VGTPLS 728
Db 532 FVLLGETMGQVTEKLQLTMYMEETCRHYVAHVKKRRVQDVISPIVPEAAYSLSHVTVGBER 591
Qy 729 AFGNLRPVL-----AEDAQRLLFTALFFPEKNCNDNICQDDLSITFSFM--SLD-----CL 777
Db 592 ELPELTPLVRWKKGQKIAQKNQTV---FERNCRSED-CAADLQLOGKLLSSMDKLTLYL 647
Qy 778 VVGSPREFNVTVTVRNDGSDSYRTQVTFPPDLDSLYKRVSTLQNRQSRWSRLACESASS 837
Db 648 ALGAVKINISINISINLGGDAYDANVSFNVSRLEFF-----INWQKEEMGISCELLES 701
Qy 838 TEVSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNK 897
Db 702 DFL-----KCSVGFPFMRKSKYEFVIFDTSHLS--GEEVLSFIVTTAQSG--NTER 750
Qy 898 TE-----FOLELPVKYAVYVMTVS-----HGVSTKYNLFTASENTS---RVMOHQY 940
Db 751 SESLHDNTLVLMVPLMHEVDTSITGIMSPTSFFVYGESVDAANFIQLDLECHFPQINITL 810
Qy 941 QVSNLQORSIPISLVFL-VPVRLN-----QTVMDRPQVTSSENLSSTCHTKER 988
Db 811 QVNTGTSTLPGSSVSISFPNRLSSGGAEMFHVQMVVGQSKGNCFSQKNPTPIIIOEQ 870
Qy 989 LPSHSDFLAELURKA-----PVVNCISIAVCORIQCDIPFGFQIEEFNATLKGNSLSD 1039

Db 371 ENIFHTIFAFYKSGRKVLDCBKGIGISCLTAHCN-----PSALAKESRTI-----D 917
Qy 1040 WYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTETKVEP----- 1084
Db 918 IY-----MLNT-EILKXSSVIO-----FMSRAKVVDPALRVEIAHGNPEEV 962
Qy 1085 ----FEVFN---PLPLIVG-----SSVGGLLLLALITAAALYKLGFPKQYKMM 1126
Db 963 TVVFEALHNLPRGVYVGVGIIAISLLVGLILPLLALLVLLWKMGFFRRYKEII 1015

RESULT 15
T31437
Integrin alpha chain SU2 - sea urchin (lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.4%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 2.3e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

Qy 336 QEGFSAATSNCP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS-----DMNDAYLG 389
Db 181 QAGFSGIIFSDNSALVMGAPGSYYLQGGYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239
Qy 390 YAAAT--ILNRNVSILVLCAPRYQHI-GLVAMFRONTGWESNANVKGTGICAYEGASLC 446
Db 240 YSLALGDFNGDGVQYVVGTPRAESLMGLVAIFDQNLQFN---QVMGTQIVAYFGYSVT 296
Qy 447 SYVDVNSGSTDVLICAPHYEQTRGGQVSVCPFPRGORARMCQDAVLYGEO----- 498
Db 297 VVDI--NNDYDILLVGPWMDGPAIQ-----RWEAGAVVYLLQNPDPVGPQA 343
Qy 499 -----GQWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRCAGVYLFHGTSGS 546
Db 344 SNRLSLSSTLIGGQIRSRPGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVVIYHG-SAN 402
Qy 547 GISPSHSQRIAGSKLS-PRLOYFGOSLGGQDLTWDGLVDLTVGAQ--CHVLLLRSPVL 603
Db 403 GLKSTPAQVLTSTLGHSGITTFGFSLOGQDMDKKYPDLLVGAESANTAVLIIRPVV 462
Qy 604 RVKATMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDRLREGIOQSV----- 658
Db 463 SLDATLNTPE-----IGINLENKTYE-LADGTWVTSFIAMT 497
Qy 659 -----VYDIALDSG-RPHSRVFNENKSTRQTVLGL-TQTCETLKLQ 702
Db 498 CFTYTGNYLPDHDIDISYTVVDSGIIANRRAMFVNDMNSIEITKRRRLAVSTQFCPLRAY 557
Qy 703 LPNCIEDPVPVILNLSLVTPLSAFN-----LRPLAEDAQRLEFALFPFEK 753
Db 558 VGNSEDKLTPIKVTLOYDL-----NNDESRLQPHIILPIDMATMSTQTKQVSIQN 609
Qy 754 NCGNDNICODLSITFSFMSLDCLVVGGRFENVTVTVRNDCGEDSYRTQVTFPFLDLSY 813
Db 610 NCVN-NICLPDLDVTV--PNLPNIVIGTQELTLDVSNRRGEDAFQSLSVYPLGLQF 667
Qy 814 RKVSTLQNRQSRWLACESASTEVSGALKSTSCSINHPIFPE-----SEVTFNIT 867

Db 668 VRL-----ERKANMDFSVTCSESD-----LRITCDTGNPMVGKNILEFGLTJLSTFQVS 717
Qy 868 FDVDS-----KASLGNKLLKLLKANVTSENNMPTNKTEFQLELPVKYAVYVMTSHGVSTK 922
Db 718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDCCTLKLLSASYPIVMYSTQED 775
Qy 923 YL-----NPTASENTRVMQHOYQV-----SNLQGRSLPIS-----LVFLVPV 960
Db 776 YVVPFPFPAKNAEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPQKNEDGEYLFYLLGI 834
Qy 961 RLNQTVIWDPRQ-----VTFSENLSSTCHTKERLPSHSDFLAELRKAPVV 1005
Db 835 MTEGVTQCQLTGKANPEGVKLEPSTRAKUSNSTTQVSGRRKRREPEVAEALQTDN--VI 892
Qy 1006 NCSIAVCQRIQCDIPFFGQIEFNAT-----LKNLSFDWYIKTSHNHLIVSTAEILF 1059
Db 893 YCASDSCVLINCTI-----DEINASKSKVVRILGRF---W-----ERTF 928
Qy 1060 NDSVFTLLPGQAFVRSQTETKVE--PFEVNP-----LP----- 1092
Db 929 QKAVSELTPVQVQATIASASAANKTIPYNIPLPRDFSDSTKASTLVTTTELVPVPVPIAW 988
Qy 1093 -LIVGSSVGGLLLLALITAAALYKLGFPKR-----QYKDMMS--EGGPP 1132
Db 989 WIIIVSVLGGIILLIILGLWKCGFPERKPKGBEKEYAPVASADKOGPP 1038

Search completed: November 25, 2003, 14:22:00
Job time : 19.9579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.17418 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFQ.....FKRQYKDMSEGPGGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5836.5	99.2	1152	1 ITAM_HUMAN	P11215 homo sapien
2	4460	75.8	1153	1 ITAM_MOUSE	P05555 mus musculus
3	3459	58.8	1163	1 ITAX_MOUSE	P20702 homo sapien
4	3401	57.8	1162	1 ITAD_HUMAN	Q13349 homo sapien
5	1547.5	26.3	1170	1 ITAL_HUMAN	P20701 homo sapien
6	1527.5	26.0	1163	1 ITAL_MOUSE	P24063 mus musculus
7	1148.5	19.5	1167	1 ITAE_MOUSE	Q60677 mus musculus
8	1140	19.4	1179	1 ITAE_HUMAN	P38570 homo sapien
9	1093.5	18.6	1151	1 ITAL_HUMAN	P56199 homo sapien
10	1084.5	18.4	1189	1 ITAH_HUMAN	Q9ukx5 homo sapien
11	1071	18.2	1170	1 ITA2_BOVIN	P53710 bos taurus
12	1057	18.0	1178	1 ITA2_MOUSE	Q62469 mus musculus
13	1055	17.9	1180	1 ITA1_RAT	P18614 rattus norv
14	1054	17.9	1181	1 ITA2_HUMAN	P17301 homo sapien
15	1051.5	17.9	1167	1 ITAG_HUMAN	O75578 homo sapien
16	663	11.3	1039	1 ITA4_MOUSE	Q00651 mus musculus
17	642	10.9	1038	1 ITA4_HUMAN	P13612 homo sapien
18	630	10.7	1035	1 ITA9_HUMAN	Q13797 homo sapien
19	593.5	10.1	1032	1 ITA4_XENLA	Q91687 xenopus lae
20	571.5	9.7	1066	1 ITA3_CRISP	P17852 cricetidae
21	567.5	9.6	1053	1 ITA3_MOUSE	Q62470 mus musculus
22	555.5	9.4	1053	1 ITA5_MOUSE	P11688 mus musculus
23	546.5	9.3	1034	1 ITAV_CHICK	P26008 gallus gall
24	545.5	9.3	1050	1 ITA5_XENLA	Q06274 xenopus lae
25	537.5	9.1	1130	1 ITA6_HUMAN	P23229 homo sapien
26	535	9.1	1044	1 ITAV_MOUSE	P43406 mus musculus
27	534.5	9.1	1072	1 ITA6_CHICK	P26007 gallus gall
28	532.5	9.0	1066	1 ITA5_HUMAN	P26006 homo sapien
29	532	9.0	1049	1 ITA5_HUMAN	P08648 homo sapien
30	526	8.9	1048	1 ITAV_HUMAN	P06756 homo sapien
31	517	8.8	1044	1 ITA8_CHICK	P26009 gallus gall
32	512.5	8.7	1091	1 ITA6_MOUSE	Q61739 mus musculus
33	498	8.5	1179	1 ITA7_MOUSE	Q61738 mus musculus

ALIGNMENTS

RESULT 1

ID	ITAM_HUMAN	STANDARD	PRT	1152 AA
AC	P11215			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor)			
GN	ITGAM OR CR3A OR CD11B			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaut M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaut M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor MO1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098993; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	494	8.4	1396	1	ITA2_DROME	P12080 drosophila
35	491.5	8.4	1146	1	ITA1_DROME	Q24247 drosophila
36	490	8.3	1033	1	ITAB_MOUSE	Q9qum0 mus musculus
37	489.5	8.3	1039	1	ITAB_HUMAN	P08514 homo sapien
38	489	8.3	1025	1	ITAB_HUMAN	P53708 homo sapien
39	486	8.3	126	1	ITAM_CAVPO	P11578 cavia porce
40	471	8.0	1181	1	ITA7_HUMAN	Q13683 homo sapien
41	470	8.0	1106	1	ITA7_RAT	P34358 rattus norv
42	462	7.9	1226	1	PAT2_CAEEL	Q63446 caenorhabdi
43	444.5	7.6	1139	1	INA1_CAEEL	Q03600 caenorhabdi
44	427	7.3	1115	1	ITA3_DROME	Q44386 drosophila
45	385	6.5	1000	1	ITA5_DROME	Q9wlm8 drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7].
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8].
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol;
RT conservation across species and homology to platelet IIb/IIIa.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9].
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10].
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11].
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tonich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment";
RL Structure 6:923-935(1998).
RN [12].
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERO DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03925; AAA59544.1; -
CC EMBL; M18044; AAA59491.1; -
CC EMBL; J04145; AAA59903.1; -
CC EMBL; S52227; AAB24821.1; -
CC EMBL; S52152; AAB24821.1; JOINED.
CC EMBL; S52153; AAB24821.1; JOINED.
CC EMBL; S52154; AAB24821.1; JOINED.
CC EMBL; S52155; AAB24821.1; JOINED.
CC EMBL; S52157; AAB24821.1; JOINED.
CC EMBL; S52159; AAB24821.1; JOINED.
CC EMBL; S52161; AAB24821.1; JOINED.
CC EMBL; S52164; AAB24821.1; JOINED.
CC EMBL; S52165; AAB24821.1; JOINED.
CC EMBL; S52167; AAB24821.1; JOINED.
CC EMBL; S52169; AAB24821.1; JOINED.
CC EMBL; S52170; AAB24821.1; JOINED.
CC EMBL; S52173; AAB24821.1; JOINED.
CC EMBL; S52174; AAB24821.1; JOINED.
CC EMBL; S52180; AAB24821.1; JOINED.
CC EMBL; S52181; AAB24821.1; JOINED.
CC EMBL; S52184; AAB24821.1; JOINED.
CC EMBL; S52189; AAB24821.1; JOINED.
CC EMBL; S52191; AAB24821.1; JOINED.
CC EMBL; S52192; AAB24821.1; JOINED.
CC EMBL; S52203; AAB24821.1; JOINED.
CC EMBL; S52212; AAB24821.1; JOINED.
CC EMBL; S52213; AAB24821.1; JOINED.
CC EMBL; S52216; AAB24821.1; JOINED.
CC EMBL; S52219; AAB24821.1; JOINED.
CC EMBL; S52220; AAB24821.1; JOINED.
CC EMBL; S52221; AAB24821.1; JOINED.
CC EMBL; S52222; AAB24821.1; JOINED.
CC EMBL; S52226; AAB24821.1; JOINED.
CC EMBL; M76724; AAA58410.1; -
CC EMBL; M84477; AAA51960.1; -
CC PIR; A31108; RWHU1B.
CC PDB; 1A8X; 17-JUN-98.
CC PDB; 1BHQ; 18-NOV-98.
CC PDB; 1BHQ; 18-NOV-98.
CC PDB; 1IDN; 25-NOV-98.
CC PDB; 1IDO; 01-AUG-96.
CC PDB; 1JLM; 11-JAN-97.
CC PDB; 1MLU; 07-AUG-02.
CC Genew; HGNC:6149; ITGAM.
CC MIM; 120980; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 4.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS00234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; 3D-structure; Repeat; Magnesium; Calcium.
CC SIGNAL 1 16
CC CHAIN 17 1152 INTEGRIN ALPHA-M.
CC DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1105 1128 POTENTIAL.
CC DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
CC REPEAT 31 84 FG-GAP 1.

Query Match 99.2%; Score 5836.5; DB 1; Length 1152;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKWVSTVMEQLKSKTLFSLMQSYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKWVSTVMEQLKSKTLFSLMQSYSEEF 196

QY 181 RHFTPEKQNNPNRSLKIPITQLLGRTHRTATGLRKVRVRELPNITNGARKNAFKILFL 240
DB 197 RHFTPEKQNNPNRSLKIPITQLLGRTHRTATGLRKVRVRELPNITNGARKNAFKILVVI 256

QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGDAFSEKSRQELNTVASKPRDHVFOIN 300
DB 257 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGDAFSEKSRQELNTVASKPRDHVFOIN 316

QY 301 NFEALXTIQNLREKIFAIEGTQTSSSSFEHMSQEGFSAATISNGPLLSVGSYDWAG 360
DB 317 NFEALXTIQNLREKIFAIEGTQTSSSSFEHMSQEGFSAATISNGPLLSVGSYDWAG 376

QY 361 GVFLYTSKSKSTFINNTRVDSNDNDAYLVYAAAILLRNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKSKSTFINNTRVDSNDNDAYLVYAAAILLRNRVQSLVLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQVSVCLP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQVSVCLP 496

QY 481 PRQARWQCDAYLYGEQGPWGRFGAALTDLGVDNGDKLTDVAIGAPGEEDNRGAVYLP 540
DB 497 PRG-RARWQCDAYLYGEQGPWGRFGAALTDLGVDNGDKLTDVAIGAPGEEDNRGAVYLP 555

QY 541 HGTSGSGISPSHSQRIAGSKSLPRLQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLLRQ 600
DB 556 HGTSGSGISPSHSQRIAGSKSLPRLQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLLRQ 615

QY 601 PVLRVKAIMEFNPREVARNVFECNDQVWKGKEAGEVRVCLHVQKSTRDLRREGIOISVVT 660
DB 616 PVLRVKAIMEFNPREVARNVFECNDQVWKGKEAGEVRVCLHVQKSTRDLRREGIOISVVT 675

QY 661 YDLALDSGRPHSRAVNETKNSTRTOTVLGLTOTCTETKLQLPNCIEDPVSPVILRLNF 720
DB 676 YDLALDSGRPHSRAVNETKNSTRTOTVLGLTOTCTETKLQLPNCIEDPVSPVILRLNF 735

QY 721 SLVGTPLSAFGLNRPVLAEDAQLFTLFPFERNKCGNDNICDDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSAFGLNRPVLAEDAQLFTLFPFERNKCGNDNICDDLSITFSFMSLDCLVVG 795

QY 781 GPREFNVTVVRNDGSDSTRVQTFPPFLDLSVRKUSTLQNRQSRQSWRLACESASSTEV 840
DB 796 GPREFNVTVVRNDGSDSTRVQTFPPFLDLSVRKUSTLQNRQSRQSWRLACESASSTEV 855

QY 841 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASIGNKLLKANVTSENMPRTNKTFF 900
DB 856 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASIGNKLLKANVTSENMPRTNKTFF 915

QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVQHQVQVSNLQQRSLPISLVLPV 960
DB 916 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVQHQVQVSNLQQRSLPISLVLPV 975

QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNVNCSIAVCQRIQCDIP 1035

QY 1021 PFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQOGAFVRSQTET 1080
DB 1036 PFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQOGAFVRSQTET 1095

QY 1081 KVEPEFVNPLPIVGVSSVGGILLALITAAALYKLGFFKRYQKQKMMSEGGPGABPQ 1137
DB 1096 KVEPEFVNPLPIVGVSSVGGILLALITAAALYKLGFFKRYQKQKMMSEGGPGABPQ 1152

RESULT 2
ITAM MOUSE
ID ITAM_MOUSE STANDARD; PRT; 1153 AA.
AC P05555;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor Mo1). ITGAM.
GN ITGAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88312584; PubMed=3044779;
RA Pytela R.;
RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";
RL EMBO J. 7:1371-1378(1988).
RN [2]
RP SEQUENCE OF 11-45 FROM N.A.
RC STRAIN-BALB/C; TISSUE=Spleen;
RX MEDLINE=86287312; PubMed=2942940;
RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;
RT "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
RN [3]
RP SEQUENCE OF 17-28.
RX MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3. THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

DR	EMB1; M29486; AAA51620.1; JOINED.
DR	PfR; A36584; RWHUIC.
DR	PDB; 1N3Y; 18-FEB-03.
DR	Genew; HGNC:6152; ITGAX.
DR	MIM; 151510; -.
DR	GO; GO:0008305; C-integrin complex; TAS.
DR	GO; GO:0004895; F-cell adhesion receptor activity; TAS.
DR	GO; GO:0004872; F-receptor activity; TAS.
DR	GO; GO:0007397; P-histogenesis and organogenesis; TAS.
DR	InterPro; IPR000413; Integrin_alpha.
DR	InterPro; IPR002035; VWF_A.
DR	Pfam; PF01839; FG-GAP_3.
DR	Pfam; PF00357; integrin_A; 1.
DR	Pfam; PF00092; vwa; 1.
DR	PRINTS; PRO1185; INTEGRINA.
DR	PRINTS; PRO0453; VWFADOMAIN.
DR	SMART; SM00191; Int alpha; 5.
DR	SMART; SM00327; VWA_1
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR	PROSITE; PS02344; VWFA; 1.
DR	KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW	Signal; Magnesium; Calcium; Repeat; 3D-structure.
KW	SIGNAL
FT	CHAIN
FT	DOMAIN
FT	TRANSMEM
FT	DOMAIN
FT	REPEAT
FT	REPEAT
FT	DOMAIN
FT	REPEAT
FT	REPEAT
FT	CA_BIND
FT	CA_BIND
FT	SITE
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CONFLICT
FT	CONFLICT
FT	SEQUENCE
QY	1 FNLDTENAMTFQENARGFGSGVVLQGSRVWGAPQEIIVAANORGSLYQCDSYTGSCPEI 60
Db	20 FNLDTELTAFRVDSAGFGSDVVQYANSVVVGAPQKITRANTGGLYCGYSTGACEPI 79
QY	61 RLQVPVEAVNMISGLSLAATTSPQLLACGPTVHQTCSNTYYKGLCFLFGSNLRQQPQK 120
Db	80 GLQVPPEAVNMISGLSLASTTSPSQLLACGPTVHHECGRMXYLTGLCFLLGPT--QLTOR 137
QY	121 PFEALRGCPQESDDIAFLIDGSGSIIPHDPFRMKEWSTVMEOCLKSKTLFSLMYSSEEF 180
Db	138 LPVSROCEPQEQODIVFLIDGSGSIGSRNFATWMNFRAVISQFPSPSTQFSLMQFSNK 197

181 RIHFTKFEQNNPRSLIKIPITOLLGRTHTATGLRKVVRLEFNITNGARKNAFKILFL 240
198 QTHFTPEFRRTNPLSLASVHQLQGFYTTATQVNVHRLPHASVGARDATKILIVI 257
241 TDGEKFGDPLGYEDVPELDRGVIYVIGVDAFRSEKSRQBELNTVASKPRPDHVPQIN 300
258 TDGKKGDSLDYKDVIPMAAAGIIRVAIGVLAFAQNRNSWKLNDIASKPSQEHIFKVE 317
301 NFEALKTIONLEKIPAIETGQTGSSSEFEHMSQEGFAAITSNGPLLSSTVGSVDWAG 360
318 DFDALDKIQNLKEKIPAIETGTTSSSEFEHMSQEGFAVFTPDGPFVIGAVGFTWSG 377
361 GVFLYTSKESKSTFNMTNRVDSMDNDAVLGVAATILRNVSQSLVLAGAPRYOHLGLVAMFR 420
378 GAFLYPNMSPFTFNMNSQENVDRDSYLGSTELALWKGVQSLVLAGAPRYOHTGKAVIFT 437
421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSCPL 480
438 QVSRQWRKAEVGTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYYEOTRGQVSCPL 497
481 PRGORARWQDAVLGCGQWGRFGAALTVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540
498 PRGWR-RWWDVAVLYGSGHPWGRFGAALTVLGVNGDKLTDVVGAPGEEENRGAVYLF 556
541 HGTSGSIGSPSHSRIAGSKLSPRLOVFGOSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
557 HGVLPGISPSHRSRIAGSKLSPRLOVFGQALSGGQDLTQDGLVLDVAGRGQVLLRTR 616
601 PVLVRKAIMPEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRREGIOISVVT 660
617 PVLVGVVSMQFIPAEIPRSFAECREQVSEQLTVQSNICLYIDKRSKNLLGSRDLQSSVT 676
661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETLKLQPNCLIEDPSPVILBLNF 720
677 LDALDPCRLSPRATFOETKRSLSRVVLGLKACENFNLLPSCVEDSVTPITLRLNF 736
721 SLVGTPLSAFNLRPLAEDAOFLTALFPEKNCNDNICODDLSITFFSMJLDCLVWG 780
737 TLVGKPLAFNLRPLAALAOYFTASLPEKNCNGADHICQDNLGIFSFPGLKSLVG 796
781 GPREFNVTVVRNDEGDSYRTQVTFPDLDSYRKVSTLQNRQSRQSWRLACASSTEV 840
797 SNLELNAEVMWMDGDSYGTITTFSPAGLSYRYVAEGQKQGLRSLHLTCDSPVGV-- 854
841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
855 SQGTWSTSCRNHILIFRGAQITFLATFDVSPKAVLGDRLLLTANVSENNMPTNKTEP 914
901 QLELPVKYAVVMVTSKYSTKYNFTAS-ENTSRVMQHOYOVNLSQSLPISLVFLVP 959
915 QLELPVKYAVTVVSSHEQTKYLNFESESEKESHVAMHRYOVNLSQSLPIVSNFWVP 974
960 VRLNQTWIDRPQVTFSENLSSTCHTKERLPSHDSFLAELKAPVNCSTAVCQRIQCDI 1019
975 VELNQAEMWDEVSHPONPSLRCSSEKIAAPPDSFLAHIKNPVLDCSIAGCLRCFCDV 1034
1020 PFGIOBEFNATLKNLSFDWIKYTSNHNLLIVSTAETLNDVSTLPGQAFVRSQTE 1079
1035 PSFVSQEBLDTLKNLSFGVVRQILQKKVSVSWVAETFTDTSVYSQLPQGAEPMAQT 1094
1080 TKVPEFVNPPLVLGVSSVGLLLALITLTAALYKLGFFKQYKDMSE 1128
1095 TVLEKRVNPTPLVGVSSIGLLLLALITAVLYKGVFFKQYKEMEE 1143

RESULT 4
ITAD HUMAN STANDARD; PRT: 1162 AA.
AC Q1349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
ITGAD.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=961111956; PubMed=8777714;
RA Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
Stauton D.E., Gallatin W.M.;
"A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3."
RT J. Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
"Cloning and chromosomal localization of a novel gene-encoding a human beta 2-integrin alpha subunit."
RT J. Biol. Chem. 275:8959-8969(2000).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
Hoffman P.A., Staunton D.E., Bochner B.S.;
"Alphabeta2 integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1)."
RT J. Exp. Med. 188:2187-2191(1998).
RN [5]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
"The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1."
RT J. Immunol. 163:1984-1990(1999).
RN [6]
RP FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-BORNE PATHOGENS, PARTICULATE MATTER, AND SENESECENT ERYTHROCYTES FROM THE BLOOD.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN ATHEROSCLEROTIC PLAQUES, AND ON SPLEENIC RED PULP MACROPHAGES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSET) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; U37028; AAB38547.1; -

DR EMBL; U40274; AAB60634.1; --
 DR EMBL; U40275; AAB60635.1; --
 DR EMBL; U40276; AAB60636.1; --
 DR EMBL; U40277; AAB60637.1; --
 DR EMBL; U40278; AAB60638.1; --
 DR EMBL; U40279; AAB60639.1; --
 DR EMBL; U40278; AAB60638.1; JOINED.
 DR EMBL; AF187881; AAF62875.1; --
 DR HSSP; P11215; IABX.
 DR Genew; HGNC:6146; ITGAD.
 DR MIM; 602453; --
 DR GO; GO:0008305; C.integrin complex; TAS.
 DR GO; GO:0004895; P.cell adhesion receptor activity; TAS.
 DR GO; GO:0016337; P.cell-cell adhesion; NAS.
 DR GO; GO:0007160; P.cell-matrix adhesion; NAS.
 DR GO; GO:0006955; P.immune response; NAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA_1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWF_A; 1.
 DR InterPro; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium;
 KW Magnesium.
 FT SIGNAL 1 17
 FT CHAIN 18 1162
 FT DOMAIN 18 1100
 FT TRANSMEM 1101 1124
 FT DOMAIN 1125 1162
 FT REPEAT 32 85
 FT REPEAT 7 2
 FT DOMAIN 150 332
 FT REPEAT 350 400
 FT REPEAT 401 452
 FT REPEAT 454 516
 FT REPEAT 518 576
 FT REPEAT 581 633
 FT CA_BIND 465 473
 FT CA_BIND 530 538
 FT CA_BIND 593 601
 FT SITE 1127 1131
 FT DISULFID 67 74
 FT DISULFID 106 124
 FT DISULFID 655 710
 FT DISULFID 769 775
 FT DISULFID 846 861
 FT DISULFID 994 1018
 FT DISULFID 1023 1028
 FT CARBOHYD 59 59
 FT CARBOHYD 87 87
 FT CARBOHYD 99 99
 FT CARBOHYD 391 391
 FT CARBOHYD 691 691
 FT CARBOHYD 733 733
 FT CARBOHYD 873 873
 FT CARBOHYD 957 957
 FT CARBOHYD 1046 1046
 FT CONFLICT 500 500
 FT CONFLICT 515 518
 FT CONFLICT 825 825
 FT CONFLICT 984 984
 FT SEQUENCE 1162 AA; 126885 MW; F296A1A3545D77D CRC64;

Query Match
 Best Local Similarity 59.3%; Pred. No. 9.3e-225; Length 1162;
 Matches :669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
 1 FNLDTENAMTFOENARGFGQSVQLQGSRRVVVGPQEIIVAAHQSGSLVQCDSYSGCEPI 60

Db 18 FNLDVEEPTIFQEDAGGFGQSVVVOFGGSRLLVGPAPLEVAANQTRLVDCAAATGMCQPI 77
 Qy 61 RLQVEVAVNNMGLSLAATTPPOLACAGPTVHOTCSNTYVKGCLFGLFNSLRQQOK 120
 Db 78 PLHIREAVNNMGLTAASTNGSRLLACAGPTLHRVCGENSYSKSGCLLSGRW-EIQT 136
 Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKWVSTVMEQLKSKXTLSFMSLSEB 180
 Db 137 VPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMGFVQAVMGQFEGCTDTLFAIMOYSNLL 196
 Qy 181 RIHFTFKFQNNPNRSLIKPTOLGRTHATGLRVKVRLEFNITNGARKNAPKILFLL 240
 Db 197 KIHTFTFOTFRYSPSQSLVDPIVLQKGLTFTATGLTIVTQTLFHHKNGARKSAKKILVI 256
 Qy 241 TDGEKFGDPLGYEDVIPBLDREGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVFOIN 300
 Db 257 TDGQYKDPLEYSVIPAOKAGIIRYVIGVGHAFQGTARQELNTISSAPPQDHVFKVD 316
 Qy 301 NFEALKTIQNLREKIFAIETGTQSGSSSFHEMSQEGFSAAITSNGLPSTVSGYDWAG 360
 Db 317 NFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGFSMSG 376
 Qy 361 GVLYTSKEKSTFTNNMTRVDSMDNDAYLGXAAAIIILNRVQSLVLGAPRYOHILVAMPR 420
 Db 377 GAFLYPNNMSPFTFNNMSENVDMDRDSYLGYSTELALWKGVQNLVLGAPRYOHTGKAVIT 436
 Qy 421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYVETRGQGVSVCP 480
 Db 437 QVSRQWRKKAQVETGTOIGSYFGASLCSVDVDSNGSTDLVLGAPHYVETRGQGVSVCP 496
 Qy 481 PRGQARWQCDVAVLYGEGQPMGRFGAALTVLGVDVNGDKLTDAVIGAPGEEDNRGAVYLF 540
 Db 497 PRGQVQWQCDVAVLRGEQHPGRFGAALTVLGVDVNEKLDVAIGAPEGQENRGAVYLF 556
 Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGSLSGQDLTMDGLVDLTVAQGHVLLLRQ 600
 Db 557 HGASEGISPSHSQRIASSQLSPRLQYFGQALSGQDLTQDGLMDLAVGARGQVLLLRSL 616
 Qy 601 PVLRVKALMEFNPREVARNFECDNVVYKGEAGVRVCLHVQKSTRRLREGQVSVVT 660
 Db 617 PVLKGVAMRFPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLOQL--GDIQSSVR 674
 Qy 661 YDLALDSGRPHSRVAFNETKSTRRTQVGLGTOTCETLKLQLPNCIEDPVSPIVLRNF 720
 Db 675 FDLALDPGLTSRAIFNETKNTPLTRRTKTLGIGHICETLKULLPCDVEDVVSPIILHNF 734
 Qy 721 SLVGTPLSAFGLNRPVLAEDAQRPLTALPPEKNCNGDNI CODDLSITFSFMSLCLVVG 780
 Db 735 SLVREPIPSQNLRPVLAAGSQDLFTASLPPEKNCQDGLCEGLGVTLFSFGLQTLTVG 794
 Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRORSORSWRLACESASSTEV 840
 Db 795 SSLEUNLVTVVWNGEDSYGTVVSLLYPAGUSHRRVSGAQKQPHOSALRLACETV-PTED 853
 Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPNTRNKTEF 900
 Db 854 EG-LRSSRCSVNHPIFHSGSNGTFIVTVDVSKATLGDRLMLRASSENKASSKATP 912
 Qy 901 QLELPVKYAVVMVTSVSHGVSTKYNLF-TASNTSRVMQHVQVSNLQGRSLDISLVFLVP 959
 Db 913 QLELPVKYAVTMTISRQBEESTKYFNFAFSDSKKKEAEHRYRVNNLSQRDLAISINFWVP 972
 Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCQRIODI 1019
 Db 973 VLLNGVAVDVMMEAPSOQL--PCVSEKRPQHSOFQLTQISRPMLDSCSIADCLQRCVD 1030
 Qy 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
 Db 1031 PSFSVQEBLDTLKNLSFGVWRETLOKKVLVSVVAEITFTDSVYSQLPQGEAFRAQWE 1090
 Qy 1080 TKVEPFEVNPPLPLIVSGSVGLLILLALITAAALYKLGFFKQYKQDMME 1128

Db 1091 MLEDEVDYNAIPINGSSVGALLLALITATLYKLGFPRHYKEMLED 1139

RESULT 5

ITAL_HUMAN STANDARD; PRT; 1170 AA.

AC P20701; O43746;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1

DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)

DE (CD11a).

GN ITGAL OR CD11a.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.

RX MEDLINE=89139587; PubMed=2537322;

RA Larson R.S., Corbi A.L., Berman L., Springer T.;

RT "Primary structure of the leukocyte function-associated molecule-1

RT alpha subunit: an integrin with an embedded domain defining a protein

RT superfamily.";

RL J. Cell Biol. 108:703-712(1989).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE=99425270; PubMed=10433829;

RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,

RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,

RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,

RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;

RT "Genome duplications and other features in 12 Mb of DNA sequence from

RT human chromosome 16p and 16q";

RL Genomics 60:295-308(1999).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.

RX MEDLINE=96036067; PubMed=7479767;

RA Ou A., Leahy D.J.;

RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha

RT L beta 2) integrin.";

RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.

RX MEDLINE=96398682; PubMed=8805579;

RA Ou A., Leahy D.J.;

RT "The role of the divalent cation in the structure of the I domain

RT from the CD11a/CD18 integrin.";

RL Structure 4:931-942(1996).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.

RX MEDLINE=99425288; PubMed=10433852;

RA Kallen J., Weizenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,

RA Cottens S., Weitz-Schmidt G., Hommel U.;

RT "Structural basis for LFA-1 inhibition upon lovastatin binding to the

RT CD11a'I-domain'.";

RL J. Mol. Biol. 232:1-9(1999).

CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,

CC ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA

CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL

CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES

CC AND MONOCYTES.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L

CC ASSOCIATES WITH BETA-2.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=P20701-1; Sequence=Displayed;

CC Name=2;

CC IsoId=P20701-2; Sequence=VSP_002738;

CC Note=No experimental confirmation available;

CC

CC -!- TISSUE SPECIFICITY: LEUKOCYTES.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC -!- DATABASE: NAME=PROW; NOTE=CD guide cd11a.htm".

CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@sib-sib.ch).

CC

CC EMBL: Y00796; CAA68747.1; -

CC EMBL: AC002310; AAC31672.1; -

CC PIR: S03308; S03308.

CC PDB: 1LFA; 29-JAN-96.

CC PDB: 1ZON; 07-DEC-96.

CC PDB: 1ZOO; 07-DEC-96.

CC PDB: 1ZOP; 07-DEC-96.

CC PDB: 1CQP; 07-AUG-00.

CC PDB: 1DQ; 03-FEB-00.

CC PDB: 1MJN; 28-JAN-03.

CC PDB: 1MO9; 14-JAN-03.

CC PDB: 1MQA; 14-JAN-03.

CC Genew; HGNC:6148; ITGAL.

CC MIM; 153370; -

CC GO; GO:0008305; C:integrin complex; TAS.

CC GO; GO:0006928; P:cell motility; TAS.

CC InterPro: IPR000413; Integrin_alpha.

CC InterPro: IPR002035; VWF_A.

CC Pfam: PF01839; FG-GAP; 3.

CC Pfam: PF00357; Integrin_A; 1.

CC PRINTS; PR01185; INTEGRINA.

CC PRINTS; PR00453; VWFADOMAIN.

CC SMART; SM00191; Int_alpha; 4.

CC SMART; SM00327; VWFA; 1.

CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.

CC PROSITE; PS0234; VWFA; 1.

CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

CC Signal; 3D-structure; Magnesium; Calcium; Repeat;

CC Alternative splicing.

CC SIGNAL 1 25

CC CHAIN 26 1170 INTEGRIN ALPHA-L.

CC DOMAIN 26 1088 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM 1089 1112 POTENTIAL.

CC DOMAIN 1113 1170 CYTOPLASMIC (POTENTIAL).

CC REPEAT 42 91 FG-GAP 1.

CC REPEAT 92 149 FG-GAP 2.

CC DOMAIN 170 349 VWFA.

CC REPEAT ? ? FG-GAP 3.

CC REPEAT 401 455 FG-GAP 4.

CC REPEAT 457 516 FG-GAP 5.

CC REPEAT 518 575 FG-GAP 6.

CC REPEAT 578 630 FG-GAP 7.

CC CA BIND 468 476 POTENTIAL.

CC CA BIND 530 538 POTENTIAL.

CC CA BIND 590 598 POTENTIAL.

CC SITE 1115 1119 GEFK MOTIF.

CC DISULFID 773 80 BY SIMILARITY.

CC DISULFID 111 129 BY SIMILARITY.

CC DISULFID 653 707 BY SIMILARITY.

CC DISULFID 771 777 BY SIMILARITY.

CC DISULFID 845 861 BY SIMILARITY.

CC DISULFID 998 1013 BY SIMILARITY.

CC DISULFID 1021 1052 BY SIMILARITY.

CC CARBOHYD 65 N-LINKED (GLCNAC...) (POTENTIAL).

```

FT CARBOHYD      89      89      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     188     188      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     649     649      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     670     670      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     726     726      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     730     730      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     862     862      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     885     885      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD     897     897      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD    1060    1060      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD    1071    1071      N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC     954     954      Q -> QGVHGLVEMOTSKOILCRPAGDAEHTVGAQEGELPC
FT                                     PWCVSEAPRDNTRAGPCR (in isoform 2).
FT                                     /FTIGVSP 002738.
FT                                     R -> W (IN REF. 1 AND 2) .
FT                                     Y -> I (IN REF. 2) .

FT CONFLICT      214      214
FT CONFLICT      660      660
FT STRAND       155      155
FT STRAND       164      164
FT STRAND       166      166
FT TURN        169      185
FT HELIX       186      188
FT STRAND      191      198
FT STRAND      202      206
FT STRAND      208      213
FT HELIX      217      221
FT TURN      222      223
FT STRAND      229      229
FT HELIX      233      243
FT TURN      244      244
FT HELIX      247      249
FT TURN      250      250
FT TURN      253      254
FT STRAND      256      263
FT HELIX      274      276
FT TURN      277      278
FT STRAND      280      286
FT HELIX      288      290
FT TURN      293      297
FT TURN      298      299
FT HELIX      300      302
FT HELIX      307      310
FT STRAND      311      314
FT TURN      318      318
FT HELIX      319      328
FT TURN      329      330
SQ SEQUENCE    1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match      26.3%; Score 1547.5; DB 1; Length 1170;
Best Local Similarity 34.3%; Pred. No. 9.le-98;
Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;

QY      1 FNLDTENAMTFO--ENARGFGOSVVOLOGSRVVVGAPQEIIVAAQNGRGSLYQCDYSTGSCE 58
DB      26 YNLDVARGARSPFPGRAGRIFGRVLQV--GNGVIVGAPGE---GNSTGSLYQCSGTGHCL 81
QY      59 PIRLOQVPEAVNMSLGLSLAATTPSPQLLACGPTVHQTCSENTYKGLCFLEGSNLR--- 115
DB      82 PVTLR--GSNYTSKYLGTWLTATPDGCSILACDPLGSRCTCDQNTYLSGLCYLFRQNLQGPW 140
QY     116 -QOPQFPALRGCPDESDIAFLIDGSGSIIPHDPRRMKEWVSTVMEOLKSKSLFSLM 174
DB     141 LQGRPGFCIRKG---NVDLVFLFDGSMLOPDEPKILDPMKMDVKKLSNTSYQFAAV 196
QY     175 QYSEEFRIHPTKPEFONNPNRSLIKPITQLGRHTATGLRVVRELENIITGARKNAF 234
DB     197 QSTSTKTFDESDYVKRQDPALLKHVHMLLTNTFGAINVATEVFEELGARPDAT 256
QY     235 KILFLITDGEKFGDPLGYEDVPELDRGVIRVIGVGDFAFRSEKSRQELNTVASKPPRD 294
DB     257 KVLIIITDGE--ATDSGNIDAAKD-----IIRYIIGIKGHFQTKESQETLHKFEASKPAGE 309
QY     295 HVFQINNFEALKTIONQLEKIFAETGTQSGSSSFEHMSQEGFSAATISNGPLSTVSG 354

```

```

DB      310 FVKILDTPEKLDLFTLELQKKIYVIEGTSKODLTSFNMELSSSGISADLSRGAHVVGAVG 369
QY     355 SYDWAGGVF--LYTSKEKSTFINMTRVSDMDMDAYLGAAA--IILNRVQSLVGLGAPRYQH 412
DB     370 AKDWAGGFLDKADLQDDTFIFIGNEPLTPEVRAGYLGVTWTWLPSPRQKSTLLASGAPRYQH 429
QY     413 IGLVAMFR--QNTGWESNANVKTQICAYFAGSLCSVDVDSNGSTDLVLIGAPHYHQT 470
DB     430 MGRVLLFOEPQGGHWSQVQIHTGTQIGSYFEGELCGVDVDQGETELLIGAPLYGEQ 489
QY     471 RGGQVSVCPLRGQRARWQCDV--LYGEOQGPWRCFCAALTIVLGVNDGDKLTDVAICAP 528
DB     490 RGRGVFIY----QRRQLGFEVSELOQDGPVPLGRFGEAITALTIDINGDLVDVAVGAP 544
QY     529 GEEDNRGAVYLFHGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTV 588
DB     545 LEE--QGAVYIFNGRHG--GLSPQPSQRIEGTQVLSGIQWFGRSIHGVNDLEGDLAVAV 601
QY     589 GAQGHVLLLRSQPVLRKVAIMEFNPREVARNVFCNDQV--KGKAGEVRVCLHVQKSTR 647
DB     602 GAESQMIVLSSRPVDMVMTLMSFSPAETPVHEVECSYSTSNMKKEGVNITICFQI--KSLY 660
QY     648 DLRREGQIQSVVTVYDLALDSGRPHSRVAFNETKSTRTQVGLTQTCTELKQLPNCI 707
DB     661 PQF--QGRIVANLTYTLQDGHRTRRGLFPGGRHELNRNIAT--TSMCTDPSFPPVCV 718
QY     708 EDPVSPIVLRNLSL---VGTPLS--AFGN-----LRPVLAEADAQRLFTALFPFPEKNCN 757
DB     719 QDLISPINVLSNLSLWEEBEGTPDQRAQKQDIPPIRLPSLHSETWEI-----PPEKNCGE 773
QY     758 DNICQDDLSITFSFMSLDCLVVGGPREFNVTYVRNODGEDSVRTQVTFPPFLDLRYKVS 817
DB     774 DKKEANLRVSFSPARSRALRLTAPASLSBELSLNLEEDAYVWQLDHFPPGLSFRKVE 833
QY     818 TLQORSORSWRLACES--ASSTEVSGALKSTCSINHPIRPENSEVTFNITFDVDSKAS 875
DB     834 ML---KPHSQIPVSCSELPEESRLLSRAL---SCNVSPIFKAGHSVALQMMFNTLVNS 887
QY     876 LGNKLKLLKANVTSENN---MPRTNKTEFQLELPVKYAVVWVTVSHGVSTKYLNFTASEN 931
DB     888 WGDVSELHANVTENNEDSDLEDNSATTI---IPILYPINILIQDESDTLVSTFKGP 944
QY     932 TSRVNHQHOYQV---SNLQGRSLP--ISLVLPVRLNQTVIWDPRPOVTFSENLSSTCHTK- 986
DB     945 KIHQVKHMYQVRIQPSIHDHNIPTLEAVVGVPPQPESEGPITHOMSVQMEPPV--PCHVED 1002
QY     987 -ERLPESHSD--FLAELRKAPVUNCIAVCQRIQCDIPFGIQEENFATLKNLSFDWYIK 1043
DB     1003 LERLPDAAEPCLPGALFPCPVW-----PQREILVQVIGTLELVGEIE 1044
QY     1044 TSHNHLIVSTAELIFNDSVFTLLPGQGAFAVRSTQETKVEPFEVNPPLIIVGSSVGGLL 1103
DB     1045 AS-SMFSLCSSLSISFSSKHFLYGSNASL-AQVMKVDVYVEKQMLYLYLVLSIGGULL 1102
QY     1104 LLALITAAALYKGLFPKRYKQKMMSEG-QPPGAEP 1136
DB     1103 LLLLIFVLYKVGKFKRLKKEKMEAGRVGPNIP 1136

RESULT 6
ITAL MOUSE
ID ITAL MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)
DE (CD11a).
GN ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
MEDLINE=91268576; PubMed=2051027;
Kaufmann Y., Tseng E., Springer T.A.;
"Cloning of the murine lymphocyte function-associated molecule-1
alpha-subunit and its expression in COS cells.";
J. Immunol. 147:369-374(1991).
[2]
SEQUENCE OF 24-42.
MEDLINE=95188276; PubMed=3887182;
Springer T.A., Teplow D.B., Dreyer W.J.;
"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
glycoproteins and unexpected relation to leukocyte interferon.";
Nature 314:540-542(1985).
-!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
LEUKOCYTES RECRUITMENT.
-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
ASSOCIATES WITH BETA-2.
-!- SUBCELLULAR LOCATION: Type I membrane protein.
-!- TISSUE SPECIFICITY: LEUKOCYTES.
-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
-!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
-!- SIMILARITY: Contains 1 VWFA domain.
-!- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M60778; AAA39426.1; -
PIR; I56126; I56126.
HSSP; P20701; 1LFA.
MG; MG1:96606; Itgal.
InterPro; IPR000413; Integrin_alpha.
InterPro; IPR002035; VWFA.
Pfam; PF01839; FG-GAP; 1.
Pfam; PF00357; Integrin_A; 1.
Pfam; PF00092; vwa; 1.
PRINTS; PR01185; INTEGRINA.
PRINTS; PR00453; VWFADOMAIN.
SMART; SM00191; Int_alpha; 5.
SMART; SM00327; VWA; 1.
PROSITE; PS00242; INTEGRIN_ALPHA; 1.
PROSITE; PS02334; VWFA; 1.
Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Magnesium; Calcium;
Repeat.
SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT ? ?
FT DOMAIN 148 334
FT REPEAT ? ?
FT REPEAT 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT CA_BIND 466 474
FT CA_BIND 528 536
INTEGRIN ALPHA-L.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
FG-GAP 1.
FG-GAP 2.
VWFA.
FG-GAP 3.
FG-GAP 4.
FG-GAP 5.
FG-GAP 6.
FG-GAP 7.
POTENTIAL.
POTENTIAL.

FT	CA_BIND	588	596	POTENTIAL.
FT SITE	1111	1115		GFPR MOTIF.
FT DISULFID	70	77		BY SIMILARITY.
FT DISULFID	108	126		BY SIMILARITY.
FT DISULFID	147	199		BY SIMILARITY.
FT DISULFID	651	705		BY SIMILARITY.
FT DISULFID	767	773		BY SIMILARITY.
FT DISULFID	840	856		BY SIMILARITY.
FT DISULFID	993	1009		BY SIMILARITY.
FT DISULFID	1017	1048		BY SIMILARITY.
FT CARBOHYD	86	86		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	185	185		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	270	270		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	444	444		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	668	668		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	696	696		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	724	724		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	728	728		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	776	776		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	857	857		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	880	880		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	890	890		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	899	899		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	927	927		N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1056	1056		N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE	1163 AA;	128343 MW;	A7A3078489E8232F CRC64;	

Query Match 26.0%; Score 1527.5; DB 1; Length 1163;
Best Local Similarity 34.1%; Pred. No. 2.1e-96;
Matches 398; Conservative 217; Mismatches 463; Indels 89; Gaps 36;

Qy	1	FNLDTENAMTFOENA-RGFQSVVQLQGSRRVVVGAPEIIVAAQNRGLYQCDYSTGCEP	59
Db	24	YNLDTRPTQSLAQAGRHFGYVQLQIEDG-VVVGAPGE--GDNTGLYHCRTSSEFCQP	79
Qy	60	IRLQVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSNTYVKGCLFGLFSGNLKQQPQ	119
Db	80	VSLH-GSNHTSKYLGMTLATDAAKGSLACDPLSRTCDQNTYLSGLYLFPOSLEGPML	138
Qy	120	KPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEMVSTVMBOLKSKTFLSLMOYSEE	179
Db	139	QNRPAYOECKMGKVDLVFLFDGSQLDKDDEKILFMDKVMRKJLNTSYQFAAVQFSTD	198
Qy	180	FRHFTFKFQ-QNNPNRSLIKPIITOLLGRTHATGLRKVKVRLFNITNGARKNAKILF	238
Db	199	CRTEFTFLDYVVKQKNDVLLGSVQPMFLLTNTFRATVYVAVHVFKEESGARDATKVLV	258
Qy	239	LITDGEKF--GDPLGYEDVIPLEDRGVIRYVIGVDGAFRSEKSRQELNTVASKPRDHV	296
Db	259	IITDGEASOKGNISAAHD-----ITRYIIGIKGHFVSVQKQKTLHIFASEPVEEFV	309
Qy	297	FQINFEALKTIQNLREKIFATEGTQTGSSSSFEHEMSQEGFSAATNSGPLLSTVGSY	356
Db	310	KILDTFEKLKDLFTDLORRIAYIEGTNRDLTSFNMLSSGSLADLSKGHAVVGAVGAK	369
Qy	357	DWAGGVF-LVTSKSEKSTFNMTVDSDMNDAYLGAYAAA-IILNRVQSLVGLGAPRYOHIG	414
Db	370	DWAGGFLDLREDLQAGTFVQEPLETSVDRGGYLGTVVAMWTSRSPRLAAGAPRYOHVG	429
Qy	415	LVAMFR--QNTGMESNANVKGQIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYYQTRG	472
Db	430	QVLLFQAPEAGGRWNTQKTEGTIGSYFGGELCSVDLDQDGEAEILLIGAPLFFGEQRG	489
Qy	473	GVSVCPPLRGQARWQCDVLYGEOGPWNGRGAALTVLGDVNGDKLTDVAIGAPGEED	532
Db	490	GRVFTY---QRRQSLFEMVSELOQDPGYPLGRFGAAITALTIDNGDRLTDVAIGAPLEE-	545
Qy	533	NRGAVYLFHGTSGSGISPHSHQRIAGSKLSPRIQYFGSLGSGQDLTMDGLVDLTVGAQG	592
Db	546	-QNAVYFNGKPG-GLSPQPSQRIQAQVFPFGIRWFGRSINGVKDLCGGDLADVVVGAEG	603
Qy	593	HVLLRSQPVLRVKAINMEFNPREVARNVFCNDQVWKGKAG-EVRVCLHVQKSTDRRLR	651

FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFD CRC64;

Query Match
 Best Local Similarity 28.8; Pred. No. 2e-70;
 Matches 355; Conservative 214; Mismatches 459; Indels 205; Gaps 43;

QY 1 FNLDTEA--WTFENARGFQSVVQLQGGVVVGAQPEIVAANORGS-----LYOCDY 52
 DB 20 FNMVDWAWTALQGPAPVLSLLHLDPSN-----NOTCLLVARRSSNRTAALYRCAL 74
 QY 53 STGCEPIRLQVPVEAVNMSLGLSLAATT--SPQLLAC-GPTVHQCSTYVYKGLCFL 109
 DB 75 SI-SPDEIACQ-PVEHICMPKRGYQGVTLVGNHNGVLVCVQVQARKFSLNSELGTACSL 132
 QY 110 FGSNLRQOPKFPALRG-----C-----PQE 131
 DB 133 LTPNLDLQAQAYFSDLEGFDPGAHVDSGDYCRSGKSGTGEEKSARRRRRTVEEED 192
 QY 132 DSDIAFLDGGSGIIPHDFRMKEMVSTVMEQL--KSKTFLFSLMOYSEFRHFTPEF 189
 DB 193 GTEIAIYLDGSGSGPQDFQAKNFIPTMMNFEKFCFECNFAVQVAVIOTFEDLOES 252
 QY 190 QNNPNRSLIKPIITQLLGRTHATGLRKYVRELFNITNGARKNAKILFLITDCEKFGDP 249
 DB 253 RDINASLAKVQSIQVKEVTKTASAMQHVLDNIFPSRSGRKKALKVMVLTGDIQFGDP 312
 QY 250 LGYEDVPELDRGVIRYIGVADAPRSEKSRQELNVTASKPRDRHVQFQINNFPALKTIO 309
 DB 313 LNLTTVINSRPMQGVAFVAFVIGVDRFKNNTYRELKLIASDPKEAHTFKVTNYSALDGLL 372
 QY 310 NOLREKFAIETGTQSSSFEHEMSOEGSAATISNGP-LLSTVGSYDAGGVFLY-TS 367
 DB 373 SKLQOIRVHMEGT---VCDALQYQLAQTGSAQLDQGVLLGTGAFNWSGGALLYSTQ 429
 QY 368 KEKSTFINMT-RVDS-DKNDAYLGAAAIILNRNVQSLVGLGAPRYQRIHGLVAMFRONTGM 425
 DB 430 NGRGCLFNQAKEDSRVTOYSLGYSLAVLHKAHGISVAGAPRHLRGAVFELKEDR- 488
 QY 426 WESNA---NVKGTQIAYFASLCSVDVDSNGSTDLVLIAGPHYEYQTRGGQVSCPLPR 482
 DB 489 -EEDAFVRRIEGQMGVFGVLCFVDIDMDGTTDFLLVAAPFYHIRGEGRVYVYQVPE 547
 QY 483 GQARWQCDVLYGEGQGWCFGAALTVLGDVNGDKLTVAIGAP-----GREDNRGA 536
 DB 548 -QDASFSLAHTLSGHPGLTNRFGFMAAAGDINQDKFTDVAIGAPLEGFGAGGAGSYGS 606
 QY 537 VYLFHGTSGSGISPSHSORIASGKSLSPRLQYFGOSLGGODLTMDGLVDLTVGAQGHVLL 596
 DB 607 VYIYNGHSG-GLYDPSQOIRASSVASGLHYFGMSVSGGLDFNGDGLADIIVGSRDSAVV 665
 QY 597 LRSQPVLRVKAIMEFNPREVARNFECNQVVKKEAGEVRVCLHVQKS---TRDLREG 653
 DB 666 LRSRPPVLDLTVSMFTFP-----DALPMVFIGKM--DVNLCEVDSSVVASBFLGLEM 715
 QY 654 QIQSVTVYDLALDSRPHSRVAFNETKNSRTRQTVLGLTQTC----- 696
 DB 716 FLNFTVDV-----TKORQLQCESSGCGSLRKNNGSGFLCEHFWLI 760
 QY 697 ETLKLQLPNCIEDPSPVTLRLNLSVGLTPLSAFENLR-----PVLAEQAQLFTALF--P 750
 DB 761 STEEL-----CEDCFENITIKVYE-----FQTSGRDRDYNPTL--DHYKEPSAIFQLP 809
 QY 751 FEKNCGNONI CODDLISITFSRMSLDCLVGGRFENVTVVRNDEGSYRQVTFPPFLD 810
 DB 810 YEKDCNKNVCFIAEIQLTN-ISOQELVVGVTKEVTNINISLTNSGDSYMTNMAINVRN 868
 QY 811 LSYRKVSTLQVORSORSLACESASSTEVSCALKSTSCSNHPIFFENSEVTNITFDV 870
 DB 869 LQFKKI-----QKVPSPDQCDPKPV---ASVLVMNCKIGHPL-LKRSSVNVSVTWQL 918
 QY 871 DSKASLGNKLLKANVTSENNMPTNKTEFQLELPVKVAVVWVVTSHGVSTKYNFTASE 930

DB 919 EBSVFNRADITVTISNSNEKSLARETR---SLOFRHAFIAVLSR--PSVVMYMN--TSQ 971
 QY 931 NTSRVWQHGYOVSNLQORSILPISLVFLVPVRLNQTVIDRPOVTFSENLSST-----CHT 985
 DB 972 SPSDHKEFFNVHGENLFGAVFQICVFIKLODF-----QIVRVNLTUKTDHTECTQ 1025
 QY 986 KERLPSPSHDLAELRKAPVNVCSIAVCORIQCDIPFFGIQEEFNATLKNLSFDMYIKTS 1045
 DB 1026 SQEPACGSDPVQHVKEHSHVCAI-----TSNKENVTVAAEISVG 1065
 QY 1046 HNHLLIVSTA-----EILFNDVSFTLLPGQAGFVRSQTETKVEPF-----EVPNPLPLIV 1095
 DB 1066 HTKQLLRDVSLEPILGEISFNKSLYEGNAE-----NHRKITVIFLKEEETRSLPLII 1119
 QY 1096 GSSVGGLLLLALITAAALYKLGPFKROYKDMWSE 1128
 DB 1120 GSSIGGLLVLLVIAILFRKGFKRKYQQLNLE 1152

RESULT 8
 ITAE HUMAN STANDARD; PRT; 1179 AA.
 ID ITAE HUMAN STANDARD; PRT; 1179 AA.
 AC P38570; Q9NZU9;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-8 precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).
 DE ITGAE.
 GN Homo sapiens (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
 RC TISSUE=Leukemia, and Lymphocytes;
 RX MEDLINE=941164962; PubMed=8119947;
 RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;
 RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
 RL J. Biol. Chem. 269:6016-6025 (1994).
 RN [2]
 RP REVISIONS TO 88-114.
 RA Parker C.M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 53-1179 FROM N.A.
 RC TISSUE=Fetal kidney;
 RX MEDLINE=20138496; PubMed=10673275;
 RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G., Stotlersauk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;
 RT "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";
 RL Genome Res. 10:165-173 (2000).
 RN [4]
 RP MUTAGENESIS OF ASP-109 AND PHE-316.
 RX MEDLINE=20400502; PubMed=10837471;
 RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;
 RT "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";
 RL J. Biol. Chem. 275:25652-25664 (2000).
 CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
 CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
 CC EPITHELIAL CELLS.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROM; NOTE=CD guide CD103 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L25851; AAB59359.2; -;
 CC DR EMBL; AF168787; AAF43107.1; -;
 CC DR PIR; A53213; A53213.
 CC DR HSSP; P11215; IABX.
 CC DR Genew; HGNC:6147; ITGAE.
 CC DR MIM; 604682; -;
 CC DR GO; GO:0008305; C:integrin complex; TAS.
 CC DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC DR InterPro; IPR000413; Integrin_alpha.
 CC DR InterPro; IPR020335; VWF_A.
 CC DR Pfam; PF01839; FG-GAP; 3.
 CC DR Pfam; PF00357; Integrin_A; 1.
 CC DR Pfam; PF00092; vwa; 1.
 CC DR PRINTS; PR01185; INTEGRINA.
 CC DR PRINTS; PR00453; VWFADOMAIN.
 CC DR SMART; SM00191; Int_alpha; 3.
 CC DR SMART; SM00327; VWA; 1.
 CC DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC DR PROSITE; PS0234; VWEA; 1.
 CC DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 CC Signal; Repeat; Polymorphism; Magnesium; Calcium.
 FT CHAIN 1 18
 FT CHAIN 19 1179
 FT CHAIN 19 177
 FT CHAIN 179 1177
 FT CHAIN 179 1124
 FT TRANSMEM 1125 1147
 FT DOMAIN 1148 1179
 FT REPEAT 181 198
 FT REPEAT 2 2
 FT REPEAT 2 2
 FT DOMAIN 145 199
 FT DOMAIN 200 391
 FT REPEAT 401 456
 FT REPEAT 457 506
 FT REPEAT 510 571
 FT REPEAT 573 638
 FT REPEAT 641 693
 FT REPEAT 641 693
 FT CA_BIND 522 530
 FT CA_BIND 586 594
 FT CA_BIND 654 662
 FT SITE 1150 1154
 FT DISULFID 70 79
 FT DISULFID 126 159
 FT DISULFID 706 762
 FT DISULFID 823 829
 FT DISULFID 893 907
 FT DISULFID 1008 1033
 FT DISULFID 1041 1057
 FT CARBOHYD 49 49
 FT CARBOHYD 271 271
 FT CARBOHYD 321 321
 FT CARBOHYD 444 444
 FT CARBOHYD 726 726

FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 360 360 D -> E.
 FT VARIANT 1041 1041 /FTID=VAR_008884.
 FT VARIANT 1041 1041 C -> S.
 FT MUTAGEN 208 208 /FTID=VAR_008885.
 FT MUTAGEN 316 316 D->A: LOSS OF E-CADHERIN BINDING.
 FT MUTAGEN 477 477 F->A: LOSS OF E-CADHERIN BINDING.
 FT CONFLICT 482 482 V -> I (IN REF. 3).
 FT CONFLICT 482 482 Q -> R (IN REF. 3).
 FT CONFLICT 950 950 R -> W (IN REF. 3).
 FT CONFLICT 1019 1019 A -> V (IN REF. 3).
 SQ SEQUENCE 1179 AA; 13008 MW; E558902EDF9D95E1 CRC64;
 Query Match 19.4%; Score 1140; DB 1; Length 1179;
 Best Local Similarity 28.9%; Pred. No. 7, 7e-70;
 Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;
 QY 45 GSLYQCOYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATSPQLLACGPTVHQ 95
 DB 65 GPLHRCSLVQDEILCHPVEHVPIKGRHGVTVVRSHHGVLCI-----QVLVRRP--HS 117
 QY 96 TCSENYVKGLCFLEGSNLROQPQ-----
 DB 118 LSSELT--GTCSLGLPDLRPOAQNFFDLENLLDPDARVDVTDGDCYNKEGGEDDVNTA 174
 QY 120 KPPEALRGCPQED-----SDIAPLDGSGSIIPHDFFRMKEWSTVMEQL-- 164
 DB 175 RRRALKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEE 234
 QY 165 KSKTLFSLMYSSEFRIFHTFKEFQNNPNRSLIKITQLLGRTHHTATGLRKYVRELFN 224
 DB 235 KCFECNFALVQGVQVTEFDLRDSQDVMSLARVQNIQVGSVTKTASAMQHVLDSIFT 294
 QY 225 ITNGARKNAFKILFLLTDEKEGDPGLGYEDVPELDREGVIRYVTVGVDAFSEKSRQEL 284
 DB 295 SSHGSRKASKVMVYLTGGIFEDPLNLTTVINSRPMQGVVERFAIGVGEFESKARTAREL 354
 QY 285 NTVASKPRDRHVQINNFEALKTIONQLREKIFAETGTOTGSSSSFEHEMSOEGFSAALT 344
 DB 355 NLIASDPETHAFKVTNYWALDGLSKLYNIISMEGT---VGDALHYQLAQIGFSAQL 411
 QY 345 SNGP--LLSTVGSYDWAGGVFLY--TSKEKSTFTINMTRVDSMDNDA----YLGVAAILRN 398
 DB 412 DERQVLLGAVGAFDWSGGALLYDTRRRGRFLNQTAADAAADAAQAQSYLGVAVVHLKT 471
 QY 399 RVQSLVLCAPRYQHIGLVAMFR--QNTGMWESNANV--KGTQICAYFGASLCSDVDVDSNGST 456
 DB 472 CSLSYVAGAPQYKHHG--AVFELQKEGREASFPLVLEGEQMGSYFGSELCPDVIDMDGST 529
 QY 457 DLVLIGAPHYBQTRGGQVSCVPLPRGORARWOCDAVLVGEQGPWGRFGAALTVLGDVN 516
 DB 530 DFLVLAAPFYVHGEGRVYVYRLSE--QDGSFLARILSGHGFNTRAREFGFAAANGDUS 588
 QY 517 GDLKTDVAIGAP----GEEDNR--GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFQQ 570
 DB 589 QDKLTDVAIGAPLEGFGADDGASFGSVIYNG--HWDGLSASPSQIRASTVAPGLQYFCM 647
 QY 571 SLUGGQDLTMDGLVDLTGVAQGHVLLLRSQPVLRVKAIMEFNPVARNVFCNQVVGK 630
 DB 648 SMAGGFDSGDGLADITVGTLLGQAVVFRSRPVVRLKVSMAFTPSALP-----IGF 697
 QY 631 KEAGSVRVLCHVQKSTRDRRLREGQIQSVVTVYDLADSGRPHSRVAFNFTKSTRQTQVL 690
 DB 698 NGVNVRLCFEI--SSVTTASESGLEALLNFLLDLDVVGKRRRLQCSQVRSCLGCLREWS 756
 QY 691 GLUTQCTETLKQLPN---CIEDPVSPIVLRNLSLVGTPLSAFNGRLRPLVAEDAQRLPT 746
 DB 757 SSGQLCEDL-LUMPTEGELCEDECFNSASVKVSYQL--QTPEQTDHPQILDRLYTPPAI 814

QY 747 ALPEFKNCGNDNICODDLSITSPMSLCLVVGGRREFNVTVTVNRDGEDSVYRTOVTF 806
Db 815 FOLPYEKACKNKLFCVAEQLA-TTVSQQLVGLTKELTLNLTNSGEDSTMTSMALN 873
QY 807 FPLDLSYKRVSTLQNRORSWRSLACESASSTEVSGALKSTCSINHPIFPENSEVTNFI 866
Db 874 YPNLQ-----LKRQKQSPSPNIQCDQPPV---ASVLIMNCRIGHVPL-KRSSAHVS 923
QY 867 TFDVDSKASLGNKLLKANVTSENN-----MPRTNKTEFQ---LELPKYAVYVWVTS 919
Db 924 VQOLEENAFNPRTADITVTVTNSNERRSLANETHLTQFRHGFVAVLSKPSIMVYVNTG 983
QY 920 S--TKYLNFTASENTSRVCHQOVSNLQORSPLISLFLVPLVRLNQTIVMDRPOVTF 977
Db 984 SHKKEFLFHVHGEN---LFGREYQ-----LQICVPTKLGLQVAVKLIRTRQ 1028
QY 978 NLSTCHTKERLPSHSDFLAELRKAPVYVNGSIAVCORIQCDIPFGIQEENFATLKNLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQVHEWHSVSCVIA-----SDKENVTVAEIS 1073
QY 1038 PDWIKTSHNHLIVST-----AEILFNDVSFTLLPGQAGFVRSQTEKVEPEVPNPL 1091
Db 1074 WD-----HSEELLKQVTELQILGEISFNKSLYEGLNAENH---RTKITVYVFLKDEKYHSL 1125
QY 1092 PLVGSVSGLLILALITAAALKLGFKKROYKDMWSE 1128
Db 1126 PIITKSGVGLLVILVILVLFKCGFFRKRYQQUNLE 1162
RESULT 9
ID -ITAL HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit.";
RL J. Biol. Chem. 268:2989-2996 (1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm".
CC PIR; A45226; A45226.
DR PDB; 1QC5; 17-MAY-00.
DR Genew; HGNC:6134; ITGAL.
DR MIM; 192968; -.
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; NAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.

DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure.
KW DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1137 1151 POTENTIAL.
FT DOMAIN 16 75 FG-GAP 1.
FT REPEAT 2 360 FG-GAP 2.
FT DOMAIN 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GPFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC... (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;
Query Match 18.6%; Score 1093.5; DB 1; Length 1151;
Best Local Similarity 27.8%; Pred. No. 1.le-66;
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;
QY 1 FNLDTENAMTFQENARG-FQGSVYQL---QGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGS 56
Db 1 FNVDVKNMTFSGPVEDMFGYTVQYVNEEGKWLIGSLVPGQPKRTGSDVYKCPVGRGE 60
QY 57 CEP-IRLOVPEA-----VNMSLGLSLAATSPPLLACGPTVHOTCSENTYVYKGL 106
Db 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTNPNNGGFLACGPLYAYRCGHLHYTTGI 119

QY 107 CFLPGNLRQOPKFFPEALRGCCQEDSDIAFLIDGSGSIIPHDFFRMKEWVSTVMEQLKK 166
DB 120 CSDVSPTFQVNSIAP--VQECSTQ-LDIVILVLDGNSIYPWDS-----VTAFLNLDLX 170
QY 167 -----SKTLFSLMOWSEEFRIHFTFKEFQNNPNPSLIKPIQTOLLGR--THRTATGLRKVV 219
DB 171 RMDIGPKQTQGVIVQIGENWTHFNFNKYSSTSEVLVAAKKIVQRGQRTMTALGDTAT 230
QY 220 RELFNITNGARKNAFKILFLTDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPR--- 276
DB 231 KEAFTEARGARRGVKKVMVITVDGESH-DNHLRLKKVIQDCEDENIQRFSAIILGSYNRGN 289
QY 277 --SEKSRQELNVTASPPDRHVPQINNFEALKTIONLREKIPAIKGTGTGSSSPEHEM 334
DB 290 LSTKEVEEIKSIASEPTEKHPFNVDDELALVTIVKTLGERIPALBATAQDAAASPEMEM 349
QY 335 SOEGFSAATTSNGPLLSLVGSDWAGGVFLYSKE-----KSTF-INMTRVDSDMNDAYL 388
DB 350 SQTGFSAHYSQDWMLGAVGAYDNGVTVMQKASIIIPRNTTFNVESTKONEPL-ASYL 408
QY 389 GYAAAIILNRVQSL-VLGAAPRYOHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCS 447
DB 409 GYTVNSATSSGVDLVYIAGOPRYNHTGQVIYRMEDGNIKILQTLSEGIQSYFGSILTT 468
QY 448 VVDNSNGSTDLVLIGAPHY-----YEQTR-GGQVSVCPPLPQGRARWQCD 491
DB 469 TDIDKDSNTDILLVGAAPMTYGTKEKRGQKVYVVALNQTRFEVQMSLEPIKQTCSSRQHN 528
QY 492 AVLYGQGGOPWG-RFGAALTVLGDVNGDKLTVAICAPGEENRGAVLPHGTSGSISP 550
DB 529 SCTTENKNEPCGARGFATAAAXDLNDLGFNDIVIGAPLEDHGGAVIYHG-SGKTIRK 587
QY 551 SHSQRITAGSKLSPRLFQSGSLGGQDLTWDGLVLTGVAQGHVLLLRQPVLRVKAIME 610
DB 588 EVAQRIPSGDGKTLAFFFQSIHNGDNLGDLTDTVIGLGAALFWRSDVAVVKVTWN 647
QY 611 FNPPEVARNVFECNDQVWKGKAG--EVRVCLHVQ-KSTRDLRREGQIOSVVTYDLALS 667
DB 648 FEPNKNVIQKNCH--MEGKETVCINATVCEVKLSKEDTIYEADLQ---YRVTLDS 700
QY 668 GRPHSAVENET-----KNSTRQOVGLGTCTETKLQLPNCIEDVPSIVLRLNFS 721
DB 701 LRQISRSFFSGTQERKQVRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDPN 752
QY 722 LVGTPLSAFNLRLPVLAEADQRLFTALFPFEKNCNDNICQDDLSITTFMSLDCLVWGG 781
DB 753 LT-DPENG-----PVLDDSLPNSVHEYIIPAKDCGKKEKISDLSHVATTEKDLIVRS 806
QY 782 PRE-FNVTVTRVNDGSDSYRTQVTFPPLDLSYRKVSTIQNQRSQSRWSRLACESSTEV 840
DB 807 QNDKFNVSILTVKNTKDSAYNTRTIVHYSNLPVSPGIEAIQKO-----SCSEN---- 853
QY 841 SGALKSTSCSINHPIFEPENSEVTFNITFDVDSKASLGN-KLLLKANVTSENWPRNKTE 899
DB 854 ----HNITKGVGPFPLRGEMVTFKILFOFNITSYLMENVTIVLSATSDSEEPETLSDNV 909
QY 900 FOLELPKAVYVMTSHGVSTKYLNFTASNTSRVMQHOYOVSN-----LQORS---- 949
DB 910 VNISIPKVEVGLQFYS-SASEVHISIAANETVPEVINSTEDIGNFINFYLRKSGSPF 968
QY 950 ---LPISLVF-----LVPVRLNQTVINDRPQVTFSENLSSTCHTKE----- 987
DB 969 MPKLKLSIFPNNTSGNPVLYPTGLSS-----SENANCRPHIFEDPFSINSK 1017
QY 988 RLPSSHDFLAELRKAPVNGSINAVCQIQCDDIPFFGQIE-----EFN 1029
DB 1018 KMTSTD---HLKRGITLDCNTCKFATITCNLTSSDISQVNSLILWKPTFKISYFSSLN 1074
QY 1030 ATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQGAFAVRSQETKVEPPEVN 1089
DB 1075 LTIIRGEL-----RENASLVUSSN-----QKRELAIOISKDGLPG 1110
QY 1090 PLPL--IVGSSVGLLLALLITALYALKYGLFFKRYKDMWSE 1128

DB 1111 RVPLWVILLSAPAGLLMLLLILALWKITGFFKRPULKQMEK 1151
RESULT 10
ITAH HUMAN
ID ITAH HUMAN STANDARD; PRT; 1189 AA.
AC Q9UKX5; Q9UKQ1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-11 precursor.
GN ITG11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=99417678; PubMed=10486209;
RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
RA Wang S.-X., Morris C.M., Krissansen G.W.;
RT "Cloning, sequence analysis, and chromosomal localization of the novel
RT human integrin alpha11 subunit (ITG11).";
RL Genomics 60:179-187(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal muscle, and Uterus;
RX MEDLINE=99395147; PubMed=10464311;
RA Velling T., Kusche-Gulberg M., Sejersen T., Gullberg D.;
RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)
RT integrin. A collagen-binding, i domain-containing, beta(1)-associated
RT integrin alpha-chain present in muscle tissues.";
RL J. Biol. Chem. 274:25735-25742(1999).
RN [3]
RP SEQUENCE OF 954-1188 FROM N.A.
RC TISSUE=Fibroblast;
RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
CC FETAL MUSCLE CELLS (IN VITRO).
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF109681; AAF01258.1; -;
DR EMBL; AF137378; AAD51919.2; -;
DR EMBL; AL359064; CAB94392.1; -;
DR HSSP; P17301; 1A0X
DR Genew; HGNC:6136; ITG11.
DR MIM; 604789; -;
DR GO; GO:0008305; C:integrin complex; TAS.

DR GO: GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO: GO:0005518; F:collagen binding activity; TAS.
DR GO: GO:0007160; P:cell-matrix adhesion; TAS.
DR GO: GO:0007517; P:muscle development; TAS.
DR InterPro: IPR000413; Intergrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR01185; INTEGRIN.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; WVA; 1.
DR PROSITE: PS00242; INTEGRIN ALPHA.
DR PROSITE: PS0234; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium; Polymorphism.
FT SIGNAL 1 22
FT CHAIN 23 1189
FT DOMAIN 23 1142
FT TRANSMEM 1143 1165
FT DOMAIN 1166 1189
FT REPEAT 38 94
FT REPEAT 102 163
FT REPEAT 167 345
FT REPEAT 359 420
FT REPEAT 422 475
FT REPEAT 477 537
FT REPEAT 539 598
FT REPEAT 601 653
FT DOMAIN 1154 1162
FT DOMAIN 1174 1177
FT CA_BIND 488 496
FT CA_BIND 551 559
FT CA_BIND 613 621
FT DISULFID 76 83
FT DISULFID 121 139
FT DISULFID 129 159
FT DISULFID 659 668
FT DISULFID 674 729
FT DISULFID 781 787
FT DISULFID 881 893
FT CARBOHYD 82 82
FT CARBOHYD 95 95
FT CARBOHYD 291 291
FT CARBOHYD 331 331
FT CARBOHYD 358 358
FT CARBOHYD 449 449
FT CARBOHYD 462 462
FT CARBOHYD 528 528
FT CARBOHYD 642 642
FT CARBOHYD 694 694
FT CARBOHYD 857 857
FT CARBOHYD 894 894
FT CARBOHYD 973 973
FT CARBOHYD 1032 1032
FT CARBOHYD 1040 1040
FT VARIANT 433 433
FT VARIANT 524 524
FT VARIANT 972 972
FT VARIANT 1003 1003
FT VARIANT 1030 1030
FT VARIANT 1094 1094
SQ SEQUENCE 1189 AA; 133609 MW; 6030308AA4CD52 CRC64;

Query Match 18.4%; Score 1084.5; DB 1; Length 1189;
Best Local Similarity 28.2%; Pred. No. 4.9e-66;
Matches 349; Conservative 210; Mismatches 500; Indels 179; Gaps 48;

QY 1 FNLDTENAMTFOENARG-FGQSVVQ--LOGSR-VVVGAPQEIIVAAQNRGLSQCDVSTGS 56
DB 23 FNDIRKPRVIFGSRTAFFGYVQQHDISGNKWLVVGAPLETNGYKQTKGDDVYKCPVIHGN 82
QY 57 CEPRL-----QVPVEAVNMSLGLSLAATTSPQOLLACGPTVHTCSENTVYKGLCLFLF 110
DB 83 CTKLNLRVTLGNVSRKDNMRLGLSLATPNKDNSFLACSPLSWSECGSSYTTGNCRSRV 142
QY 111 GSNLRQOPQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKWVST---VMEOLKK- 166
DB 143 NSNFRSKTAP-ALQRC-QTYMDIVLDGNSIYP-----WVEQHFILNLLKXF 192
QY 167 ----SKTLFSLMQYSEEPRIHFTFKFQNNPNRSLIKPITQLLG-RTHATATGLRKVRE 221
DB 193 YIGPGQIQGVVQYGEDVVHFEHLNDYRSVKDWEAASHIEQGGTETRTAFGIEPARSE 252
QY 222 LFNITNGARKNAFKILFLTDEKFGDPLGYEDVIPELDEGIRVIRVIGV-----GDAPR 276
DB 253 AFQ--KGRKRGAKVMIVITDGESHDP--DLEKVIQOSRDNVTRYAVAVLGYNNRRGIN 309
QY 277 SEKSRQELNTVASKPRDRHVFOINNFEALKTIONQLREKIFAIEGTOTGTSSSSFEHMSQ 336
DB 310 PETFLNEIKYIASDDPKHFFNVTDAAALKDIVDALGDRIFSLGKTNK-NETSGLEMQQ 368
QY 337 EGFSAAITSNGLPLSTVGSYDMAGVFLYTSKEK-----STFINMTRVDSMDNDAYLGYA 391
DB 369 TGFSSHVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPELKNHGAFLVYT 428
QY 392 AAILLRNV-QSLVLGAPRYQHIGLVAMP-RONTGMWESNANVKGTOIGAYFGASLCSDV 449
DB 429 VTSWSSRQGRVVVAGAPRNFHTGKVFLETHMNRSLTIHQAMRGOQIGSYFGSEITSD 488
QY 450 VDSNGSTDLVLGAPHYEQTR-GGOVSVCPLPRGORARWQCDVAVLGGQPMWGFAGAA 508
DB 489 IDGGVTDVLLVGAPMYFNEGRGKVYTEL---KQNRFPVNGTLKDSHSYONARFGSS 545
QY 509 LTVLGVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYF 568
DB 546 IASVRDLNQDSYNDVVVGAPLEDNHAGAIYIFHGFRGS-ILKTPKQRIASELATGLQYF 604
QY 569 GQSLSGGQDLTMGLVDLTGVAQGHVLLRSQPLVRKAIEMFNPREVARNVF--ECNDQ 626
DB 605 GCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQINASLHPEPSKI--NIHRDC--- 659
QY 627 VVKGEAGEVRVCL-----HVQKSTRDLREGQIQSVVYTDLALDAGPHSR 673
DB 660 ----KSGRDTCLAAFLCETPIFLAPHFQTTVG-----IRYNATMDERRYTTPR 705
QY 674 AVFNETKNS-TRQTOVLGTLTCTETKLQLPNCIEDPVSPIVLRNLSLVGTPLSAFGN 732
DB 706 AHLDEGDRFTNRVALLSSQELCERINFHVLDTADYVYKPTVFSVEYSLEDP-----D 758
QY 733 LRPVLADAORLTALFPPEKNGCNONI CODDL----- 765
DB 759 HGPMLDDGWPTTLRVSVPFMWNGCNEDEHCVPLDLVLDARSDLPTAMEYCORVLRKPAODCS 818
QY 766 SITFSPMSLDCLVVGGPREFNVTVRNDCGDSYRQVTFPPFLDLRYKRVSTLQORSQ 825
DB 819 AYTLSDFTTVFIIESTRQRVAVBEATLENRGENAYSTVLNIGSANLOF--ASLIQKEDSD 876
QY 826 RSNRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLIKAN 885
DB 877 GS--IECVNEER-----RLQKQVCNVSYPFFRAKAKVAFRLDDEF-SKISFLHLEIELA 928
QY 886 VTENNMPRTNKTKE--FQLELPVKYAVVMVVTSHGVSTKY---LNFTAS--ENTSRVMQH 938
DB 929 AGSDSNRSDTKEDNVAFLRFLHLYEADVLFTRSSLSLHYEVKLNSSLERYDGTGPPFSC 988
QY 939 QYQVSNLGO---RSLPISLVLPVRLNQTVIWRDPOVTFSENLSSTFC----HTKRLPS 991
DB 989 IFRIQLNLPFIHIGIMMKITIPATRSNGNLLKLRLDFT-DEVANTSCNIGWNSTEYRPT 1047

Db 370 VGFSAEYSPONNIMLGAVCADWSGTGVQKTPHGLHIFSKQAEQILQDRNHSYLGYS 429
Qy 392 AAILNRVOSVLGAPRYQHIGLVAMFRONTGWESNANV-----KGTQICAYGASL 445
Db 430 VASISTGNSVHFVAGAPRANVTQGVLYSYN-----ENGNTVTIQSGRQDQIGSYFGLV 484
Qy 446 CSVDVDSNGSTDLVLGAPHYEQTR--GGQVSCPLPRQGRARWQCDVAVLYGEOQPMW 503
Db 485 CAVDVNKDTITDVLGAPMYMDLKEGRVFLFTITKG-ILNWH--QFLEGPNGLENA 541
Qy 504 RFGAALTVLGVDNGDKLTVAIGAAGEDNRGAVYLFHGTSGSGISPHSHQRIAGS--KL 561
Db 542 RFGSAIAALSDINDGPNVDIVGSPLENQNSGAVIYVNGHEGM-IRLYSQKILGSDRAF 600
Qy 562 SPRLQYFGSLSGQDITMGLVDLTVGAQHVLLLRSPQVLVAKMEFNPREVARNVF 621
Db 601 SSSLQYFGRSLDGYDGLNGSDITDVSVAFGQVQVQLMSQSIADVDSATPTPKI--TLL 658
Qy 622 ECNDQVYKGEAGVRVCLHVQKSTRDRLRGQIQSVVTVYDLALD----SGRPHSRAVEN 677
Db 659 NKNAEI-----KLKLCF-----SAKRPNTQNNQVAIVYNTIDEDQFSSRVISGLFK 707
Qy 678 ETKNSTRQTOVLGLTQCB--TLKQLPNCIEDPVSPIVLRNFSL--VGTPLSAPGNL 733
Db 708 ENNERCLQTMIVSQAKRCSEYIIHQEPS---DIISPLMLCMNISLENPGT-----756
Qy 734 RPVLAEADQRLFTALFPKXKCGNDITCODLSITF---SPMSLDCVLVGGPREFNVTV 789
Db 757 NPALAEYSETVKVFSIPFKDCGDDGVCISDLVLNVQOLPATQQPPIVSNQMKRLTFVS 816
Qy 790 TVRNDGDSYRTOVTFFPFLDLSYRKVSTLQNRQSRQSWRLACESASST-EVSGALKSTS 848
Db 817 QLKWKESAYNTEIUVDFSENLF-----ASWSPVDGTEVTQIASQSKSVT 864
Qy 849 CSINHPIPPENSEVNTFTDVSQKSLGNKLLKANVTSENMPRTNKTQFQLELPVKY 908
Db 865 CNVGPALKSKQQTFTTFDFNLQ-NLQNASISFRALSQBENMADNSVNLKLSILY 923
Qy 909 AVYVMTSHGVSTKYLNFATSENTRVMQHQYQVSNLQOR-----SLPISLFLV 958
Db 924 DAEIHT-RSTNINIFYSLDGNVSVV-HSFE--DIGPKFIFSIVKVTGSPVPSMA---976
Qy 959 PVRLNQTVMIDRPOVTFSEN--LSSTCHTKE-----RLPSSHDFLAE- 998
Db 977 -----SVIIHPQVTKDKNPLMYLTGVHTDQAGDISCEAINPLKIGTSSSVSFKSEN 1030
Qy 999 LRKAPVNVCSIAVCQRIQCDIPFFGQIEEFNATLKNLSFDWYIKTSHNHLIIVSTAEI- 1057
Db 1031 FRHIKELNCRATSCSINCMWLRDLQVKEGYFLNVSTRIWNGTFAASTFQTVQLTAAAEID 1090
Qy 1058 LFNSVFTL-----LPGQAFVRSQETKVEPFE-VPNPLPLIVGSSVGLLILALITA 1110
Db 1091 TYNQIYVIEENTVTIP-----LTIMKPEKVEVPTGVIVGSIAGIILLALLVA 1140
Qy 1111 ALYKLGFPKRYKDM 1125
Db 1141 ILWKLGFPRKRYEKM 1155

RESULT 12
ITA2 MOUSE
ID ITA2_MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)
DE (collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologous supports collagen and laminin adhesion but
RT not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z29987; CA82877.1; --
CC EMBL; X75427; CA53178.1; --
CC PIR; S41142; S44142.
CC HSP; P17301; IAOX.
CC MGD; MGI:96600; Itga2.
CC InterPro; IPR00413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 4.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Platelet; Signal; Repeat; Calcium; Magnesium.
CC SIGNAL 1 26
CC CHAIN 27 1178
CC DOMAIN 27 1129
CC TRANSMEM 1130 1151
CC DOMAIN 1152 1178
CC REPEAT 42 100
CC REPEAT ? ?
CC DOMAIN 185 375
CC REPEAT ? ?
CC REPEAT 431 483
CC REPEAT 485 546
CC REPEAT 548 607
CC REPEAT 612 664
CC CA_BIND 496 504
CC POTENTIAL.

FT CA_BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 GFPR MOTIF.
FT DISULFID 80 89 BY SIMILARITY.
FT DISULFID 677 734 BY SIMILARITY.
FT DISULFID 786 792 BY SIMILARITY.
FT DISULFID 862 873 BY SIMILARITY.
FT DISULFID 1016 1047 BY SIMILARITY.
FT DISULFID 1052 1057 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 696 696 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;
Query Match 18.0%; DB 1; Length 1178;
Best Local Similarity 27.8%; Pred. No. 3.7e-64;
Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;
QY 1 FNLDENAMTFQ--ENARGFQSYVOL---QGRVVVVGAPQEIIVAAQNRGLSYQC--DYST 54
DB 27 YNGLPGAKIFSGPSSEQFGYVQQLTNPQGNWLLVGSWSPGPPENRMDGVYKCPVDLPT 86
QY 55 GSCEPIRLQ-----VPVAVNMSGLSLAATSPQQLACGPTVHQTCSNTYVKGILC 107
DB 87 ATCEKLNQNSASISNVITKTNMSLGLTLTRNPGTGGLTCGPLWAHQGNQYVATGIC 146
QY 108 FLFGSNLRQPOQ---KFPALRGCPQEDSDIAFLIDGSGIIPHPFRMKWVSTWEOAL 164
DB 147 ----SDVSDPDLTFSFAVQACPSL--VDVVVVCDESNIYP--WEAVNFKLVFVTGL 199
QY 165 K--KSKTLFLSMQYSEFRIHFTKFEQNNPRSLIKPITQLLG--RTHATGLRKVRRE 221
DB 200 DIGPKKTQVALIQYANEPRIIFNLNDFETKMDVQATSETRQHGGLTNTFRAIEFARDY 259
QY 222 LFNITNGARKNAPKILFLTDGKFGDPLGYEDVIPELDRGVIRVIGV-----GDAFR 276
DB 260 AYSQTSGRPGATKVMVVTGDGSH--DGSKLKTVIOCCNDDELTRFGIAGVAGLYLNRLAD 318
QY 277 SEKSRELANTVASKPRDRHVQINNFPEALTKIQNLQREKIFATEGTQTGSSSFHEMSQ 336
DB 319 TKNLKEIKAIATPTPTERYFFNFVADAEALKEKAGTLGEQIFSTEGTVQG--GDNFQEMAQ 377
QY 337 EGFSA--AITSNGPLSLTVSGYDWAGGVFLYTSKESKSTFINMT--RVDSDMN--DAYLGYA 391
DB 378 VGFSAADYAPQNDILMLGAGFADWSGTLVQETSHKPVIFPKQAFDQVLDQRNHSFLGYS 437
QY 392 AAILRNVRQSLVGLAPRYQHIGLVAMFRONTGMESNANV---KGTQIGAYFGASLCS 447
DB 438 VAAISTEDGVHVFAGAPRANYTCQIVLYSNK---QGNVTVIQSHRGGDQIGSYFGSVLCS 494
QY 448 VDVDSNGSTDVLIGAPHYVEQTR--GGGVSVCLPRGORARWOCDAVLVGGQGPWGRF 505
DB 495 VDVDDKTIITDVLVGLAPTNDLKKEEGVYFTITKGLNHOQ----FLEGPEGTGNARF 551
QY 506 GAALTVLGVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSORIASGKLSPR- 564
DB 552 GSAIALSDINMGDFNDVIGSVENSGAVIYNGHGT--IRTKYSQKILGNSGAFRR 610
QY 565 -LQYFGQSLGGQDLTMDGLDVTGAAQGHVLLRSQPVLRKVAIMEFNPVARNVFE 623
DB 611 HLQFFCRSLDGVGLNGSDITDVSIGALGVQLWQSIADVAIEALFTP----- 660
QY 624 NDQVVKGEAGEVRVCLHVQKSTRDLREGQTSVVTYDLAD-----SGRPHSRVAFNET 679
DB 661 -DKITLNLKDAKITLKLCPRAEFAPAGQNNQV--AILFNMTLADAGHSRVTSRGVFREN 717

QY 680 KNSTRQTOVLGLTQTCTET--LKQLPNCIEDPVPSPVILRNFLSVLGTPLSFAFGNLRPVL 737
DB 718 SERFLQKMMVNEVQKCEHHSIQKPS---DVNPLDLRVDISLENPGTS-----PAL 768
QY 738 AEDAQRLLFTALPPFKKNGNDNICODDLIS1-----TFSEMSLDCLVVGGPREFNV 788
DB 769 EAYSETVKVFISPIPFYKCEGSDGICUSDLILDVQQLPAIQTSF-----IVSNQNKRLTFS 823
QY 789 VTVNRDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACASST--EVSGLKST 847
DB 824 VILKRGESAYNTVVLAEFSENLF-----ASFSPVDGTETVCEVSSQKSV 871
QY 848 SCSINHP1PPENSEVTFNITFDVSKASLGNKLLKANVTSENMPRTNKTE--FQLELP 905
DB 872 TCDVGPALKSEQQVTFTFINFDNLQ--NLQAAINFOAFSESO--ETNKADNSVSLTIP 928
QY 906 VKYAVYVMVTSHGVSSTKYNLFTASNTSRVMQHVQVSNLGR-----SLPISLV 955
DB 929 LLYDAELHLT--RSTNINFEISSDENAPSVIK---SVEDIGPKFISLKVTAGSAPVMA 984
QY 956 FLV-----PVRLNQTVMIDRPOVTF--SENLS 980
DB 985 LVTHIPIQYTKENPLLYLTG1QTDQAGDISCTAEINPLKLPHTA-----PSVSFKENFR 1040
QY 981 STCHTKERLPKSHSDFLAELRKAPVNVVNCIAVCQRIQCDDIPFGIQEENFATLKNLSFDM 1040
DB 1041 ---HTKE-----LDCRTTSCNITCWLKDLHMAEYFINVTRVMNRT 1080
QY 1041 YIKTSHNLLIVSTAEILFNDSVFTLLPQCGAFVRSQETKVEPEVENPPLVIGSSVG 1100
DB 1081 FAASTFQVQVTAABIDTHNPQLPFIENAVTIPLMIMKPTKEAEVPT--GVIIGSTIA 1138
QY 1101 GILLALITAAALYKLGFFKQYKDM 1125
DB 1139 GILLALLANTAGLWKLGFKQYKDM 1163
RESULT 13
ID ITAL RAT
AC P18614; STANDARD; PRT; 1180 AA.
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
DE (CD49a).
GN ITGAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RX MEDLINE=90338125; PubMed=2380249;
RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
RA Esch F., Carbonetto S., Reichardt L.F.;
RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor
for laminin and collagen."
RL J. Cell Biol. 111:709-720(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
RX MEDLINE=99313197; PubMed=10386626;
RA Nolte M., Pepinsky R.B., Venyaminov S.Y., Koteliansky V.,
RA Gotwals P.J., Karpusas M.;
RT "Crystal structure of the alphabeta integrin I-domain: insights into
integrin I-domain function."
RL FEBS Lett. 452:379-385(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

Db 877 CESN-----ONITCRVGYPLRAGETVTFKIIQFNTHSLSENAIHLSATSDSBE 927
QY 891 NMPTNTKTEQLEPLVKYAV---YMWVTSHGVS-----KYLNTASENTSRVQHQ 939
Db 928 PLESNDNEVNIPIKVEYGLQFYSSASEHHSVAANETIPEFINST--EDIGNEINV 985
QY 940 YQVSNIGQSRSLP---ISLVP-----LVPRLNQTVMD-----RP----- 971
Db 986 YTIKRGHFPMPQLQISFPNLTADGYPVLYPIG-----WSSSDNVNCRPSRSLDPFG 1039
QY 972 -----QVTS-----ENLSSTCHTKERLPSPSHDSFLAELRKAPVNCISAVCORIQ 1016
Db 1040 INSGKMTKSKSVLRKGTQDCSSTC-----GVATITCSLLPSDLQ 1082
QY 1017 CDT-----PFFGIOBEF---NATLKGNSLFDWYIKTSHNHLIVSTABILFNDSVFTLL 1067
Db 1083 VNVSLLLWKEPTF-IRAFSSNLTLRGELK-----SENSLTSSN----- 1123
QY 1068 POGAFVRQTEKVPFEVNPPLPL--IVGSSVGLLLALITAAALYKLGFFKQYKDM 1125
Db 1124 -----RKRELATQISKDGLGRVPLWILLSAFAGLLLLMLLILALWKGFFKRLPKKK 1177
QY 1126 MSE 1128
Db 1178 MEX 1180

RESULT 14
IT22_HUMAN STANDARD; PRT; 1181 AA.
AC P17301;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
RC TISSUE=Endothelial cells;
RX MEDLINE=89308879; PubMed=2545729;
RA Takada Y., Hemler M.E.;
RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit
RT (platelet GPIa): homology to other integrins and the presence of a
RT possible collagen-binding domain.";
RL J. Cell Biol. 109:397-407(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
RX MEDLINE=98019223; PubMed=935332;
RA Embley J., King S.L., Bergelson J.M., Liddington R.C.;
RT "Crystal structure of the I domain from integrin alpha2beta1.";
RL J. Biol. Chem. 272:28512-28517(1997).
RN [4]
RP VARIANT HPA-5 (BR).
RX MEDLINE=94043762; PubMed=7901236;
RA Santosco S., Kalb R., Walka M., Kiefel V., Mueller-Eckhardt C.,
RA Newman P.J.;
RT "The human platelet alloantigens Br(a) and Br(b) are associated with a
RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
RT alpha 2)." ;
RL J. Clin. Invest. 92:2427-2432(1993).
RN [5]
RP VARIANT GLU-534.
RX MEDLINE=20206009; PubMed=10744142;

RA Kroll H., Gardemann A., Fechter A., Haberboesch W., Santosco S.;
RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
RT gene polymorphism on coronary artery disease and acute myocardial
RT infarction.";
RL Thromb. Haemost. 83:392-396(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
CC associates with beta-1. Interacts with HPS5.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- POLYMORPHISM: POSITION 534 IS ASSOCIATED WITH PLATELET-SPECIFIC
CC ALLOANTIGEN HPA-5 (BR). HPA-5A/BR(A) HAS LYS-534 AND HPA-5B/BR(B)
CC HAS GLU-534. HPA-5B IS INVOLVED IN NEONATAL ALLOIMMUNE
CC THROMBOCYTOPENIA (NAIT OR NATP). THE K534E POLYMORPHISM MAY PLAY A
CC ROLE IN CORONARY ARTERY DISEASE (CAD).
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17033; CAA34894.1; -;
DR EMBL; AF512556; AAM34795.1; -;
DR PIR; A33998; A33998
DR PDB; 1A0X; 25-NOV-98.
DR PDB; 1DZ1; 02-AUG-01.
DR Genew; HGNC:6137; ITGA2.
DR MIM; 192974; -;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007596; F:blood coagulation; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 1181 INTEGRIN ALPHA-2.
FT DOMAIN 30 1132 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1133 1154 POTENTIAL.
FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1155 1161 INTERACTION WITH HPS5.
FT REPEAT 45 103 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 188 378 VWFA.

FT REPEAT 378 433
FT REPEAT 434 486
FT REPEAT 488 549
FT REPEAT 551 610
FT REPEAT 615 667
FT CA_BIND 507
FT CA_BIND 563 571
FT CA_BIND 627 635
FT SITE 1157 1161
FT DISULFID 83 92
FT DISULFID 680 737
FT DISULFID 789 795
FT DISULFID 865 876
FT DISULFID 1019 1050
FT DISULFID 1055 1060
FT CARBOHYD 105 105
FT CARBOHYD 112 112
FT CARBOHYD 343 343
FT CARBOHYD 432 432
FT CARBOHYD 460 460
FT CARBOHYD 475 475
FT CARBOHYD 699 699
FT CARBOHYD 1057 1057
FT CARBOHYD 1074 1074
FT CARBOHYD 1081 1081
FT VARIANT 534 534

TURN 170 171

STRAND 173 180

TURN 183 184

HELIX 188 199

TURN 200 201

STRAND 204 204

TURN 206 207

STRAND 209 216

STRAND 220 224

TURN 226 228

HELIX 232 240

TURN 241 241

HELIX 252 262

TURN 263 264

HELIX 266 268

TURN 269 269

STRAND 275 282

HELIX 289 291

TURN 292 301

STRAND 302 303

STRAND 304 311

HELIX 313 317

TURN 318 319

HELIX 323 330

TURN 331 332

HELIX 337 340

STRAND 341 344

HELIX 347 353

HELIX 354 362

TURN 363 363

SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match

Best Local Similarity 17.9%; Score 1054; DB 1; Length 1181;

Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;

Qy 1 FNLTENAMTQ-ENARGFQSVVQL---QGSVVVVGAPQEIIVANQSGSYQC--DYST 54

Db 30 YNVLPEAKIFSGPSEQFGYAVQGFNPKGNWLLVGSPPSGFPENRMGDYKCPVDLST 89

Qy 55 GSCPIRLQ-----VPVEAVNMSLGLSLAATSPPOLACGPTVHTQCSNTYVKGIC 107

Db 90 ATCEKLNQSTSTPNVTEMTKNMSLGLTRNMGTTGGFLTCGPLMWAQQCNQYTTGVC 149

Qy 108 FLFGNLRQQPKPFALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMYEQLK-- 165

Db 150 SDISPDF-QLSASFSPATQCPSL-IDVVVVCDSENSIYPWD--AVKNFLEKFEVQGLDIG 205

Qy 166 KSKTFLSLMQYSEEFRIHETTFEPQNNPNRSLIKITQLLG-RHTATGLRKKVRELN 224

Db 206 PTKQTQGLIOYANNPRVFNLTNYTKKEEMI VATSQTSQGGDLTNTFGAIOYARYAYS 265

Qy 225 ITNGARKNAFKILLLTLDGKFGDPLGVEDVPELDEGVIRYVIGV-----GDAFRSEK 279

Db 266 AASGGRSRSATKVMVVVTDGESH-DGSMKXAVIDQCHNDILREGIAVLGVLPNRALDTKN 324

Qy 280 SROELMTVASKPRDRHVQINNFEALKTIONQUREKIPAEIGTQTGSSSSFEHEMQEFG 339

Db 325 LIKEIKAIASIPTEYFFNVNDEAALLEKAGTLGEQIFSEGTGQVQ-GDNFQEMQVGP 383

Qy 340 SAAITSNP--LLSTVGSYDMAGGVFLYTSKESKSTFINMT--RVSDMN-DAYLGYAAAI 394

Db 384 SADYSSQNDILMUGAFGWSGTIVQKTSHGHLIFPKOAFDQILQDRHSSYLGYSVAA 443

Qy 395 ILNRVQSLVLGAPRYQHIGLVAMPFRQNTGMESNANV-----KGTQIGAYEGASLCV 448

Db 444 ISTGESTHFVAGAPRANYTQIVLYSVN-----ENGNITVIAHRGDQIGSYFGSLCV 498

Qy 449 DVDSNGSTDLVLIGAPHYYEQR--GGQSVVCLPRQGRARMOCDVLYGEEQOQPNRFG 506

Db 499 DVDKDTITDVLVAGAPMYMSDLKKEGRVYVFTIKKILGQHQ---FLEGEGIENTRFG 555

Qy 507 AALTVLGDVNGDKLTDVAIGAPCEENRGAVLFGHTSGSISPSHSQRTAGS--KLSPR 564

Db 556 SAJAAALSDINMGDFNDVI VGSPLENQNSGAVIYNGHQGT-IRTKYSQKILSGDGAFRSH 614

Qy 565 LOYFGOSLSGGQDLTMDGLVDLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNPECN 624

Db 615 LOYFGRSLDGYDLNGDSITDVSIGAFQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 672

Qy 625 QVVKGEAGEVRVCLHVQKSTRDLRREGIOISVTVYDLALD---SGRPHSRVAFNETK 680

Db 673 AQII-----LKLCE---SAKFRPTKQNNQVAVVYNTLTDADGFSRVRTSGLPKENN 721

Qy 681 NSTRQTOVLGLTQTC--ETLKLQLPNCIEDPSPVLRNLFNLSLVGTPLSAFNLPRVLA 738

Db 722 ERCLQNMVNVQAQSCPEHIIYIQEPS---DVVNSLDLRVDISLENPGTS-----PALE 772

Qy 739 EDAQLRFTALFPPEKNCNDNICQDDLSITF---SPMSLDCLVWGPREFNVTVVRND 794

Db 773 AYSETAKVFSIPHKDCGEDGLCISDLVDVRQIPAAQEQPFIVSNQNKELTFSVTLKNK 832

Qy 795 GEDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWRLACESASST-EVSGALKSTSCSINH 853

Db 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTEVTCQVAAQSAQSVACDVGY 880

Qy 854 PIPPESENTEFNITFDVDSKASLGNKLLKANVTSENNMPTNKTETQLELPVKYAYMV 913

Db 881 PALKRQGVTFITNFDNLQ-NLQNASLSFQALSEQENKADNLVNLKIPLYDAEI- 938

Qy 914 VTSHGVSITYKLNFTASENTSRVNHQVQVSNLQOR-----SLPISLVFLV----- 958

Db 939 ---HLTRSTNINFEISSDGNVPSIVHSFEDVGPKFIFSLKVTGTVSVPMATVITHI PQ 995

Qy 959 -----PVRNLQTVIMDRPOVTF-SENLSSTCHTKER 988

Db 996 YTKENKPLMYLTGVQTDKAGDISCNADINPLKIQGT-----SSSVSPKSNFR---HTKE- 1047

Qy 989 LPSHSDFLAELRKAPVNVNCISAVCQRIQCDIPFGLOEENATLKNLSFDWIKTSHNH 1048

Db 1048 -----LNCRATSCSNVTCMLKDVHMKGEYFVNVTTRINGTASSTFQT 1091

Qy 1049 LLIVSTAEEI-LFNDVSFTLLPGOGAFVRSQETETKVEPFEPNP-----LP--LIVGSSVG 1100

Db 1092 VQLTAAAEINTYNPEIYVI-----EDNTVTIPLMIMKDEKAEVPTGVIIGSIIA 1141

Qy 1101 GLLLALLITAALYKLGFFKQYKDM 1125

Db 1142 GILLIALVALWKLGFRRKRYEM 1166

RESULT 15

ITAG HUMAN STANDARD; PRT; 1167 AA.

AC 075578; Q9UH28;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-10 precursor.

GN ITGA10

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Articular chondrocytes;

RX MEDLINE=98352078; PubMed=9685391;

RA Camper L., Hellman U., Lundgren-Aakerlund E.;

RT "Isolation, cloning, and sequence analysis of the integrin subunit

RT alpha10, a beta1-associated collagen binding integrin expressed on

RT chondrocytes.";

RL J. Biol. Chem. 273:20383-20389 (1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Endothelial cells, and Heart;

RX MEDLINE=20169197; PubMed=10702680;

RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,

RA Wang S.-X., Langley R., Krissansen G.W.;

RT "The integrin alpha10 subunit: expression pattern, partial gene

RT structure, and chromosomal localization.";

RL Cytogenet. Cell Genet. 87:238-244 (1999).

CC - FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.

CC - SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10

CC ASSOCIATES WITH BETA-1.

CC - SUBCELLULAR LOCATION: Type I membrane protein.

CC - TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSIONS IN

CC MUSCLE AND HEART. FOUND IN ARTICULAR CARTILAGE.

CC - DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC - SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC - SIMILARITY: Contains 1 VWFA domain.

CC - SIMILARITY: Contains 7 FG-GAP repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR ENBL; AF074015; AAC31952.1;

DR ENBL; AF112345; AAF21944.1;

DR ENBL; AF172723; AAF61638.1;

DR HSP; P17301; IAOX

DR Genew; HGNC:6135; ITGA10.

DR MIM; 604042; -

DR GO; GO:0008305; C: integrin complex; TAS.

DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.

DR GO; GO:0005518; F: collagen binding activity; TAS.

DR GO; GO:0007160; P: cell-matrix adhesion; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR02035; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.

DR SMART; SM00327; vwa; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.

DR PROSITE; PS0234; VWFA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

KW Signal; Repeat; Calcium; Magnesium.

FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 1167 INTEGRIN ALPHA-10.

FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1123 1145 POTENTIAL.

FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).

FT REPEAT 38 97 FG-GAP 1.

FT REPEAT 7 7 FG-GAP 2.

FT DOMAIN 167 350 VWFA.

FT REPEAT 365 427 FG-GAP 3.

FT REPEAT 428 482 FG-GAP 4.

FT REPEAT 483 545 FG-GAP 5.

FT REPEAT 546 605 FG-GAP 6.

FT REPEAT 608 660 FG-GAP 7.

FT DOMAIN 1134 1140 POLY-LEU.

FT CA_BIND 494 502 POTENTIAL.

FT CA_BIND 558 566 POTENTIAL.

FT CA_BIND 620 628 POTENTIAL.

FT DISULFID 76 86 BY SIMILARITY.

FT DISULFID 666 675 BY SIMILARITY.

FT DISULFID 681 736 BY SIMILARITY.

FT DISULFID 789 795 BY SIMILARITY.

FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 844 844 I -> L (IN REF. 2).

FT CONFLICT 909 909 G -> V (IN REF. 2).

FT CONFLICT 926 926 E -> D (IN REF. 2).

SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;

Query Match 17.9%; Score 1051.5; DB 1; Length 1167;

Best Local Similarity 28.7%; Pred. No. 8.7e-64;

Matches 351; Conservative 199; Mismatches 491; Indels 183; Gaps 45;

QY 1 FNLDTENAMTFQENARG-FGQSVVOLQSGR-----VVVGAPEQETVAANQSGLYOC----- 50

DB 23 FNLDDEHPRLFGPPEAEFGYVLQHVGGQRMVLVGAPEWDGSGRRGDVTRCPVGGAH 82

QY 51 -----DYSTG-SCEPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTQCS 99

DB 83 NAPCAKAGHLGDYQLGNSSHP-----AVNMHLGMSLLETDCDGGFMACAPLWSRACGS 134

QY 100 NTYVVKGLCFGLGSNLRLQOPKPEALRGCPQSDSDIAFLDGSGLIIPHDFRMKRW--V 157

DB 135 SVFSSGICARVDASFPQSGSLAPTAQR-CPTY-MDVVIVLDGNSNIYP-----WSEV 184

QY 158 STVMEOL-----KSKTLFLSMQVSEEFRIHFTKFEQNNPNRSLIKPITOLLGR-TH 210

DB 185 QIFLRLVGLKFLIDPEIQIQLVQVGSFVHWSLGDFTKVEVVAANKLSRREGRETK 244

QY 211 TATGLRKVVRELFNITNGARKNAFKILFLTLTGKEF-GDPLGYEDYVPELDREGVIRYVI 269

DB 245 TAAQIMVACTEGFSQSHGGRPEARLLVVVTDSHSDGEEL--PAALKACEAGRVTRYGI 302

QY 270 GV-GDAFRSEKS-----RQELNTVASKPPRDHFQINFEALKTIQNQLREKIPAIGTQT 324

DB 303 AVLGHYLRQRDPSSFLREIRTIASDPDERFFFNVTDEAALTDIVDALGDRIFGLEGSHA 362

QY 325 GSSSPFHEMSQEGFAAITSNGPLLSITVSGVDWAGGVFLYTSKEKSTFINTRVDS--- 381

DB 363 ENESSFGLMSQIGFSTHRLKKGILFGMVGYDMGWSVLWLEGHRLFPFRMALEDFPP 422

QY 382 --DMNDAYLGYA-AAIILNRVQSLVGLGAPRYQHIGLVAMFR-QNTGMWESNANVKGTOI 437

Db 423 ALONHAAYLGYSVSMLLRGRRFLSGAPRFRHRGKVIAFOLKKDCAVRVAQSLQEQI 482
Qy 438 GAYEGASICSVDVDSNGSTDVLIGAPHYY--EOTRGQVSVCPPLRGQARWQCDVLY 495
Db 483 GSYFGSELCPDTRDGTDDVLLVAAPMFGLPQNKETGRVYVYLV--GQOSLLTLQGTLO 540
Qy 496 GEOQPMGRFCAALTVLGDVNGDKLTVAIAPGEEEDNGAVLFGHTSGSGISPSHSQR 555
Db 541 PEPQD-ARFGFANGALPDNLNODGFADVAVGAPLEDGHOGALYLYHGTQ-SGYRPHPAQR 598
Qy 556 IAGSKLSPRLOYFQOSLSGGODLTMDGLVDLTGAAQGHVLLLRQPVLRVKAIMEFNPRE 615
Db 599 IAAASMPHALSYFGRSVDGRDLDDGLVDVAVGAQGAAILLSRPVHLTPSLEVTPQA 658
Qy 616 VARNVFCNDQVVKGEAG--EVRVCLHVQKSTRDLREGQIOSVVTYDIALDSGRPHSR 673
Db 659 ISVVQDCRR--RGQAVCLTAALCFQVTSRTFGRWDH---QPYMRFTASLDEWTAGAR 712
Qy 674 AVFNET--KNSSTRQTOVLGTLTOTCETLKLQPCIEDPVPSPVILRLNFSLVGTPLSAFG 731
Db 713 AAFDGSQORLSPRLRLSVG-NVTCQQLHFHVL-DTDLRPPVALTVTFALDNTTKPG-- 768
Qy 732 NLRVLAEDAQRFTALFPPEKNCNGNDNICODDLSITFSFMSLDC-----LVGGPR 783
Db 769 ---PVLNEGSPTSIQKLVPSKDCGPDNECVTDLVQ---VNMDIRGSRKAPFVVRGRR 822
Qy 784 EFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEVSQA 843
Db 823 KVLVSTTLNKENAYNTSLIIFSRNL---HLASLTPQR-ESPIKVECAPSA----- 872
Qy 844 LKSTSCSINHPIFPENSEVTENITFDVDSKASLG---NKLL-----LKANVTSENNMPT 895
Db 873 -HARLCSVGHVPFQTGAKVTLLEFEPSCSLLSQVFGKLTASSDSLSLRNGTLQENTAQT 931
Qy 896 NKTEFQLELPVKYAVYVWVTSHGVS TKYLNFTASENTSRVMOHY----- 940
Db 932 -----SAYIOYEPH-----LLFSESTLHRYEVHPYGTLPVGPGEFKTTL 972
Qy 941 QVSNLG---QRSLSPLSLVLP-----VRLNQTVIDRPOVTFSENLSSTCHTKER 988
Db 973 RVQNLGCVVSGLIISA--LLPAVAHGGNYFLSLSQVI-----TNNASCIVQNLTE 1021
Qy 989 LPSHSDFLAELKAPVNVNCSTAVCQRIQCDIPFFGIQEEFNATLKNLSFDWYIKTSHNH 1048
Db 1022 PPGPPVHPEELOHTNRLNGSNTQCVVRCHLGQLAGTEVSVGLLRVHNEFFRRAKFS 1081
Qy 1049 LLIVSTAEILFNDSVFTLLPGGAFVRSQTEKVEPVPNPLPIVGS SVGGLLLLALI 1108
Db 1082 LTVVSTFELGTEGSLQLTEASRWSESLLEV-VQTRPILISLWILIGSVLGLLLALL 1140
Qy 1109 TAALYKLGFF-----KROYK 1123
Db 1141 VFCLWKLGFFAHKKIPEEKREK 1164

Search completed: November 25, 2003, 14:17:30
Job time : 13.1742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:11:54 ; Search time 33.1068 Seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481b-4
Perfect score: 5884
Sequence: 1 FNLDTENAMTFQENARFGQ.....FKRQYKDMMSGCGPPGAEPO 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriapi:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4371	74.3	1151	11	Q9J130
2	3891.5	66.1	1036	11	Q8CA73
3	3799.5	64.6	920	6	Q28984
4	3470	59.0	1169	4	Q81VA6
5	3310.5	56.3	1169	11	Q9QXH4
6	3236.5	55.0	1161	11	Q9QYF7
7	1523.5	25.9	1161	11	Q9WTV4
8	1513	25.7	1160	11	Q9R200
9	1395	23.7	1196	13	Q98TF1
10	1350.5	23.0	1086	4	Q96HB1
11	1344.5	22.9	1187	13	Q98TF0
12	1269	21.6	927	6	Q8HZV0
13	1154.5	19.6	1167	11	Q88340
14	1109	18.8	1167	11	Q88341
15	1041	17.7	1171	13	O42094
16	1032	17.5	1038	11	Q8BS01

17	1005	17.1	895	11	Q9WUF8
18	998.5	17.0	1160	6	Q8MKF4
19	865	14.7	348	4	Q8TES5
20	860	14.6	1332	5	Q9BPQ8
21	796	13.5	205	11	Q63001
22	753.5	12.8	780	13	Q06271
23	738	12.5	823	4	Q8WFI8
24	686.5	11.7	823	11	Q8CE84
25	669	11.4	1032	11	Q61989
26	644.5	11.0	1033	6	Q9BGU3
27	643	10.9	1036	11	Q91YD5
28	628.5	10.7	257	11	Q8C270
29	619.5	10.5	1041	5	Q9UB90
30	614.5	10.4	1041	5	O76378
31	580	9.9	1034	13	Q98TT7
32	579.5	9.8	1054	5	Q9U6S1
33	550	9.3	1033	13	O42598
34	534	9.1	1016	13	Q91779
35	530	9.0	974	11	Q324W2
36	529	9.0	1073	11	Q8CC06
37	526	8.9	1047	6	Q9MZD6
38	525.5	8.9	1007	6	Q9GK48
39	512.5	8.7	1034	6	O9TUN4
40	509.5	8.7	1036	6	O9TUN6
41	506.5	8.6	1012	11	O70304
42	491.5	8.4	1049	5	Q8SY51
43	475	8.1	833	5	Q9BPQ7
44	469.5	8.0	1036	6	O9TU44
45	468.5	8.0	604	11	Q8BQ25

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
 ID Q9J130
 AC Q9J130
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Integrin beta 2 alpha subunit.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fathallah D.M. Sr.; Zerrila K. Jr.;
 RT "Cloning of the rat cd11b cDNA sequence."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF268593; AAF81280.1; -
 DR HSSP; P11215; 1BHQ
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 5.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; WFA; 1.
 SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.3%; Score 4371; DB 11; Length 1151;
 Best Local Similarity 72.6%; Pred. No. 9.4e-317;
 Matches 826; Conservative 150; Mismatches 159; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARFGQSVVVLQGSVVVVCAPQEIVAANQSGSLYOCDYVSGCEPI 60
 DB 17 FNLDTENPMTFQENASFGQSVIQLGETRVVVAAPQEVKAVNQTGALYQCDYSTNRCDPI 76

```
Qy 61 PLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLSGNLRRQPOK 120
Db 77 PLQVPEAVNMSLGLSLAATTVPPOLLACGPTVHONCKENTYVNGCLYFGLSGNLRRPQQ 136
Qy 121 FPEALRGCPQSDSIAFLIDSGSGSIIPHDPRMKEWSTVMEQKSKTKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQESNIAFLIDSGSGSINTIDFKMKEFVSTVMDQFKSKTKTLFSLMQYSD 196
Qy 181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVRVRELFTNGARKNAFKILFL 240
Db 197 RTHFTFNDKRNPPKSHVRPIQLNGRTKTASGIRKVRVRELFTNGARKNAFKILFL 256
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVIGVDGAFRSKSOELNTVASKPPRDHVFQIN 300
Db 257 TDGKFGDPLGYEDVPELDREGVIRVIGVDGAFRSKSOELNTVASKPPRDHVFQIN 316
Qy 301 NFPEALNTIQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 360
Db 317 NFPEALNTIRNLQLEKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFTNMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 377 GAFLYPSKDKASFINTTRIDSDMDNDAYLGASVIRNRVOSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKTQICAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGGOVSVCP 480
Db 437 QNTGMWESNANVKTQICAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGGOVSVCP 496
Qy 481 PRGQARWQCDVLYGEGQPGWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQPGWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 615
Qy 601 PVLKATMEPNPREVARNVFECDQVVKGEAGEVRVCLVQKSTRDLREGIOQSVT 660
Db 616 PVLKATMEPNPREVARNVFECDQVVKGEAGEVRVCLVQKSTRDLREGIOQSVT 675
Qy 661 YDLALDGRPHSRVFNSTKSTRQTQVLGTOTCETLKLQPCNIBDPVSPVLRINF 720
Db 676 YDLALDGRPHSRVFNSTKSTRQTQVLGTOTCETLKLQPCNIBDPVSPVLRINF 735
Qy 721 SLVGTPLSAFGLNRPVLAEDQRFTALPPEKKNCGNDNI CODDLSITFMSDCLVVG 780
Db 736 SLVGTPLSAFGLNRPVLAEDQRFTALPPEKKNCGNDNI CODDLSITFMSDCLVVG 795
Qy 781 GPRFENVTYVRNDCGDSYRQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 796 GPRFENVTYVRNDCGDSYRQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 854
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 855 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 914
Qy 901 QLELPVKYAVNMTSHGVSTKYLNTFTASENTSRVQHQYVSNLQGRSLPISLVFLVPV 960
Db 915 QLELPVKYAVNMTSHGVSTKYLNTFTASENTSRVQHQYVSNLQGRSLPISLVFLVPV 974
Qy 961 RLNQTVIWDNRQVTFSENLSCTCHTKERLPSSHDFLAELRKAPVNCISIANVORLOCDIP 1020
Db 975 RLNQTVIWDNRQVTFSENLSCTCHTKERLPSSHDFLAELRKAPVNCISIANVORLOCDIP 1034
Qy 1021 PFGIOEFNATLKGNSLFDWYIKTSHNHLIIYSTAEILFNDSVFTLLPGQAGFVRSQTE 1080
Db 1035 PFGIOEFNATLKGNSLFDWYIKTSHNHLIIYSTAEILFNDSVFTLLPGQAGFVRSQTE 1094
Qy 1081 KVEPEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGABPQ 1137
Db 1095 KVEPEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGABPQ 1151
```

```
RESULT 2
Q8CA73 PRELIMINARY; PRT: 1036 AA.
ID Q8CA73
AC Q8CA73;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Integrin alpha M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
SQ EMBL; AK039444; BAC0350.1; -.
SR SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;
```

```
Query Match 66.1%; Score 3891.5; DB 11; Length 1036;
Best Local Similarity 66.0%; Pred. No 4.8e-281;
Matches 751; Conservative 125; Mismatches 143; Indels 119; Gaps 2;
```

```
Qy 1 FNLDTENAMTFQENARFGQSVVOLGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFQENARFGQSVVOLGSGTSVVAAAPQEAQAVNQTGALYQCDYSTSRCHPI 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLSGNLRRQPOK 120
Db 77 PLQVPEAVNMSLGLSLAVSTVPQOLLACGPTVHONCKENTYVNGCLYFGLSGNLRRPQQ 136
Qy 121 FPALRCPCQEDSDIAFLIDSGSGSIIPHDPRMKEWSTVMEQKSKTKTLFSLMQYSEEF 180
Db 137 FPALRCPCQESDIIVFLIDSGSGSINNIDQKMKFVSTVMEQFKSKTKTLFSLMQYSD 196
Qy 181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVRVRELFTNGARKNAFKILFL 240
Db 197 RIHFTFNDKRNPPKSHVRPIQLNGRTKTASGIRKVRVRELFTNGARKNAFKILFL 256
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVIGVDGAFRSKSOELNTVASKPPRDHVFQIN 300
Db 257 TDGKFGDPLGYEDVPELDREGVIRVIGVDGAFRSKSOELNTVASKPPRDHVFQIN 316
Qy 301 NFPEALNTIQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 360
Db 317 NFPEALNTIQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFTNMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGASVIRNRVOSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKTQICAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGGOVSVCP 480
Db 437 ENFTWEPHTSIKG----- 450
Qy 481 PRGQARWQCDVLYGEGQPGWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 451 ----- 450
Qy 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 451 -----SQRIIGAHFSPGLQYFGQSLSGKDLTMDGLMDLAVGAQGHVLLRAQ 498
Qy 601 PVLKATMEPNPREVARNVFECDQVVKGEAGEVRVCLVQKSTRDLREGIOQSVT 660
Db 499 PVLKATMEPNPREVARNVFECDQVVKGEAGEVRVCLVQKSTRDLREGIOQSVT 558
```

```
QY 661 YDIALDSGRPHSRVAVNETKSTRROTQVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720
DB 559 YDIALDPGRSIRIRAFFDETKNTRRRRTQVGLMKQCKETLKLIPDCVDSVSPILRLNY 618
QY 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLCLVVG 780
DB 619 TLVGEPLRSFGLNRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLCLVVG 678
QY 781 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRORSRWL-ACESASSTE 839
DB 679 GPQDFNKSVTLRNDGSDSYRTQVTVYPSGLSYRKDSASQNPITKKPWFVKPAESSSSE 738
QY 840 VSGALSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANKVTSNNMPTNKTE 899
DB 739 GHGALKSTTWNHPIPPANSEVTNITFDVDSHAGSFGNKLKKAIVASENNMSTRHTK 798
QY 900 FQLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMOHOYQVSNLQORSPLSLVLP 959
DB 799 FQLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMOHOYQVSNLQORSPLSLVLP 959
QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCIAVQRIQCDI 1019
DB 859 VQINNVTVMDHPQVIFSONLSSACHTEQKSPHSNFRDQLERTPVLCNSAVACKRQCDL 918
QY 1020 PFFGIOEFNATLKGNSLSPWYIKTSHNHLIYSTAIELFNDVSFTLLPQOGAFVRSQTE 1079
DB 919 PSFNTQEIFNVLTKGNSLSPWYIKTSHGHLVLSSTELFNDVSFAFALLPQOESYVRSKTE 978
QY 1080 TKVEPEVNPPLIIVGSSVGGLLALLALITAAIYKLGFFKQYKDMWSEGGPGAPBQ 1137
DB 979 TKVEPEVNPPLIIVGSSVGGLLALLALITAGLYKLGFFKQYKDMWSEAAQDAPPQ 1036

RESULT 3
Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE CD11b (Fragment).
GN CD11b.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RP Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWF; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON_TER 1
FT NON_TER 920
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.6%; Score 3799.5; DB 6; Length 920;
Best Local Similarity 78.8%; Pred. No. 2.9e-274;
Matches 726; Conservative 84; Mismatches 110; Indels 1; Gaps 1;

QY 118 PQKPEALRCQPEDSDIAFLIDSGSIIPHDFRMKEWTVMYEQLKSKTFLSLMQYS 177
DB 1 PQKPEALRCQPEDSDIAFLIDSGSIIPHDFRMKEWTVMYEQLKSKTFLSLMQYS 60
NCBI_TaxID=9606;
```

```
QY 178 EBFRIHTFFKEQNPNPNSRLIKPTTOLLGRTHRTATGLRKVVRELFTNTNGARKNAFKIL 237
DB 61 EDFYHTFTFNDPKRNPSPKLLVRPRLQLGRTHRTATGLRKVVRELFTSHSKSGARENAFKIL 120
QY 238 FLTDTGKFGDPLGVYEDVTPELDREGVIRYVIGVDAPRSEKSRQELNTVASKPRDHPV 297
DB 121 VVITDGEKFGDPLGVYEDVTPEDARKGVIRYVIGVDADFNSKWSRELNNTIASKPCGDHPV 180
QY 298 QINNFEALKTIQNLQREKIFAIEGTQTSSTSSSFEHMSQEGFSAAITSNGLPSTVGSYD 357
DB 181 QVNNFEAVKTIQNLQREKIFAIEGTQTSSTSSSFEHMSQEGFSAAITSNGLPSTVGSYD 240
QY 358 WAGGVFLYTSKESKSTFINNTRVDSMDNDAYLYAAAAILRNVRQSLVLGAPRYQHIGLVA 417
DB 241 WAGGAFHLPKORVIFINTTRVDSMDNDAYLYAVEVILRNQAQSLVLGAPRYQHTGLV 300
QY 418 MFRONTGMWESNANVKGTQIGAVFGASLCSVDVDSNGSTDVLICAPHYVETRGGOVS 477
DB 301 MFKQNSGAWEKWADIKGSIQISYFGASLCSVDVNRDGGSDVLICAPHYVETRGGOVS 360
QY 478 CPLPRGORARMCQDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAV 537
DB 361 CPLPQG-RAKMQCRVILCGEQGHPSRFGAALTALGDVNGDKLTDVAIGAPGEEDNRGAV 419
QY 538 YLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLL 597
DB 420 YLFHGTSELGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLL 479
QY 598 RSQPVLRKAIEMFNPREVARNVFCNDQVVKGEVYVCLHVOKSTRDLRREGQIOS 657
DB 480 RSQPVLRVBAASVVFPREVARNVYECRQAAKTQIAGEVQVCLQVRKSTWDRLRREGDTS 539
QY 658 VVTDLALDSGRPHSRVAVNETKSTRROTQVGLTQTCETLKLQLPNCIEDPVSIVLR 717
DB 540 IITYDLALDPGRPHPRAVEETKNTRROTQVGLSRKCEHLALWLPDCVEDSVTVILR 599
QY 718 LNFSLVGTPLSAFGLNRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLCL 777
DB 600 LNFSLVKGKPASSFGLNRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSFMSLCL 659
QY 778 VVGGRPEFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRORSRWLACESAS 837
DB 660 VVGGRPDLKVLTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRORSRWLACESAS 719
QY 838 TEVSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANKVTSNNMPTNK 897
DB 720 TEESTALKSTSCSINHPIPPENSEVTNITFDVDPDAFLGYKLLKANKVTSNNMPTNK 779
QY 898 TEFQLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMOHOYQVSNLQORSPLSLVFL 957
DB 780 TEFQLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMOHOYQVSNLQORSPLSLVFL 839
QY 958 VVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCIAVQRIQ 1017
DB 840 VVRLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCIAVQRIQ 899
QY 1018 DIPFGIOEFNATLKGNSL 1038
DB 900 DIPFGIOEFNATLKGNSL 920

RESULT 4
Q81VA6 PRELIMINARY; PRT; 1169 AA.
AC Q81VA6;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```



```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Blood;
RA      Strausberg R.;
RL      Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL: BC038237; AAH38237.1;
SQ      SEQUENCE 1169 AA; 128521 MW; A17B484FEFC79EB6 CRC64;

Query Match      59.0%; Score 3470; DB 4; Length 1169;
Best Local Similarity 61.0%; Pred. No. 1.7e-249;
Matches 689; Conservative 142; Mismatches 292; Indels 6; Gaps 4;

QY      1  FNLDTENAMTFOENARFGOSVVOQLQGSRRVVVGAPQBIIVAAANORGSLYQCDYSTGSCPEI 60
DB      20  FNLDTEELTAFRDVDSAGFGOSVVOQYANSWVVGAPQKITAANTQGGYQCGYSTGACEPI 79
QY      61  RLQVPVEAVNMNLSGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLFGLGNSLRQOPQK 120
DB      80  GLQVPPPEAVNMNLSGLSLASTTSPOLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTOR 137
QY      121  FPEALRCGPQEDSDIAFLIDGSGSIIIPHDFRMKEWSTYMEQLKSKTFLSLMOYSEEF 180
DB      138  LPVSRQECPRQEQDIFLIDGSGSISRNFMATMNFVRAVISQFQRPSTQFSLMQFSNKF 197
QY      181  RIHTEFEFQNNPNSRLIKPITOLLGRTHATGLRKVVRELFTNGARKNAFKILFL 240
DB      198  QTHTEFEFRKSSNPLSLASVHQLQGTFTTATAIQNVVHRLFHASGARDAKILIVI 257
QY      241  TDGKFGDPLGYEDVIPELDREGVIRYVIGVDGAFRSEKSRQELNTVASKPPRDHVFQIN 300
DB      258  TDGKKEGDSLDYKDVIPMDADAAGIIRVAIGVGLAFQNRNSWKELNDIASKPSQEHFKVE 317
QY      301  NFEALKITQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB      318  DFDALKDITQNLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGPVLGAVGFSFTWSG 377
QY      361  GVFLYTSKEKSTFTNMTVRDSDMDAYLGGYAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
DB      378  GAFLYPPNMSTFTINMSQENVDMRSDYLCYSTELALWKGVSLVGLGAPRYOHTGKAVIFT 437
QY      421  QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYYEQTGQGVQVCP 480
DB      438  QVSRQWRMKAETVGTQIGSYFGASLCSVDVDSGSDTDLVLGAPHYYEQTGQGVQVCP 497
QY      481  PRGQARQOCDAVLYGEOGQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNKGAVLYF 540
DB      498  PRGWR-RWMCDAVLYGEOGHPGRFGAALTVLGVDVNGDKLTDVWIGAPGEKENRGAVLYF 556
QY      541  HCTSGSGISPSFSQRIAGSKLSPRLQYFGQSLSGGODLTMDGLVDLTVGAQGHVILLRSQ 600
DB      557  HGVLTGPSISPSHSQRIAGSKLSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLTR 616
QY      601  PVLRYKAIEMFNPVARNVFECNDQVYVKGKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
DB      617  PVLWVGVMQFIPABI PRSAFECREQVVSQTLVQSNICLYIDKRSKNLLGSLDQSSVT 676
QY      661  YDLALDSRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
DB      677  LDALDPLGRSLSPRATFQETQNRSLSRVRLGLKAHCENFNLLPLSCVEDSVTPITLRLNF 736
QY      721  SLVGTPLSAFGLNRLPVLAEADAQRLFTALFPPEKNGCNDNICODLSITFSFMSLDCLVVG 780
DB      737  TLVGLPLAFNRLPWLAAQAQRYTASLPPEKNCADHICQDNLIGISFSPGLKSLVVG 796
QY      781  GPREFNVTVVRNDEGDSYRTQVTEFFFLDLISYRKVSTLQNRQSRWRLACESASSTEV 840
DB      797  SNLELNAEVMWVNDGSDSYGTTVTFSSHAGLSYRVVAEGQKQQLRSLHLTCDSPAVG-- 854
QY      841  SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLVNTVSENMPRTNKTFF 900
DB      855  SQGTWSTSCRINHILFRGGAQITFLATFDVSPKAVLGDRLLLTANVSENNTPTSKTTF 914
QY      901  QLELPVKYAVVMVTSHGVSITKYLNFAS-ENTSRVMQHQYQVSNLQGRSLPISLVLVP 959

915  QLELPVKYAVVMVTSHEQFTKYNLNFSESEKESHVAMERYQVNNLQGRDLFVSNFWVP 974
960  VRLNQTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCORIQCDI 1019
975  VELNQEAVMVDVEVSHQPNPSURCSEKTAIPASDFLAHIQKNPVLDSCSIAGCLRFRCDV 1034
1020  PFGIOEEFNATLKGNSLSPDWIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSOTE 1079
1035  PSFSVQEEELDTLKGNSLFGWVRQILQKKVSVVVAEITFDTSVVSQLPQGEAFMAQTT 1094
1080  TKVPEPEVNPPLIVSSVGGLLLALITALYALYKLGFPKQYKQNMSE 1128
1095  TVLEKYKVHNPPLIVSSIGGILLALITAVLYKVGFFKQYKEMMEE 1143

RESULT 5
QYQXHA PRELIMINARY; PRT; 1169 AA.
AC QYQXHA;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang X., Goreki K., Tong C., Rattiss F.-M., Tseng S.-Y., Pardoll D.,
RA Teuchiya H.;
RT "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF211864; AAF23492.1; -.
DR HSP; P11215; IBIHQ.
DR MGD; MGI:96609; Itgax.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR Integrin.
KW SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match      56.3%; Score 3310.5; DB 11; Length 1169;
Best Local Similarity 56.6%; Pred. No. 1.3e-237;
Matches 645; Conservative 173; Mismatches 307; Indels 19; Gaps 7;

QY      1  FNLDTENAMTFOENARFGOSVVOQLQGSRRVVVGAPQBIIVAAANORGSLYQCDYSTGSCPEI 60
DB      20  FNLDTEELTAFRDVDSAGFGOSVVOQYANSWVVGAPQKITAANTQGGYQCGYSTGACEPI 79
QY      61  RLQVPVEAVNMNLSGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLFGLGNSLRQOPQK 120
DB      80  GLQVPPPEAVNMNLSGLSLAATTPPOLLACGPTVHHTCRENIYLTGLCFLLSSSPKQS-QN 138
QY      121  FPEALRCGPQEDSDIAFLIDGSGSIIIPHDFRMKEWSTYMEQLKSKTFLSLMOYSEEF 180
DB      139  FPTAQECPCQDQDIFLIDGSGSISSTDFEKMILDFVKAVMSQLQRPSTFSLMQFSDF 198
QY      181  RIHTEFEFQNNPNSRLIKPITOLLGRTHATGLRKVVRELFTNGARKNAFKILFL 240
DB      199  RVHTTFNFISSPLSLGSRQRLRGYTYTASAKHVITELFTTQSGARQDATKVLVI 258
QY      241  TDGKFGDPLGYEDVIPELDREGVIRYVIGVDGAFRSEKSRQELNTVASKPPRDHVFQIN 300
```

```

Db 259 TDGRKQDNLSDVSVIPMAEASIIIRYAGVGKAFYNEHSKQELKAIASMPSEHYVFSV 318
QY 301 NFEALKTIONQRLKEKIFALEGTQSSSFEHMSQEGFSAATISNGPLLSVTGSDWAG 360
Db 319 NFDALXDIENQKKEKIFALEGTETPSSTFEHMSQEGFSAVFTDGPVLGAVGFSWSG 378
QY 361 GVFLYTSKESKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 379 GAFLYPSNMRPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIFT 438
QY 421 QNTGMWESNANVKGTOIGAYFGASCLSDVDVDSNGSTDVILIGAPHYYEOTRGQVSVCLP 480
Db 439 QSRHRWPKSEVRGTQIGSYFGASCLSDVDMDRSDGLVLI GPHYYEOTRGQVSVCPM 498
QY 481 PRQARWOCDAVLYGEOQOPWGRFCAALTLDVGNVGDKLTDVAIGAPCEEDNRGAVYLF 540
Db 499 P-GVGRWCGTTLHGEQHPWGRFCAALTLDVGNVGDLSLADVAIGAPCEEDNRGAVYLF 557
QY 541 HGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 558 HGASRQDIAPSPSORISASQIPSRIOYFGQSLSGGQDLTRDGLVDLAVGSKRGLLRTR 617
QY 601 PVLRVKAIEMFPREVARNVFECNDQVVGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 618 PILRVSPVTHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIHESPKTL--GDLRSTVT 675
QY 661 YDLALDSGRPHSRVNETKNSRRROTQVLGLTQTCETIKLQLPNCIEBPPVSVILRLNF 720
Db 676 FDALDHGRLSTRAIFRETKTRALTTRVKTGLGNKHCESVKLLLPACVEDSVPTITLRLNF 735
QY 721 SLVGTPLSFAFNLPRVLAEDAORLFTALPFPFKNCGNDNICDDLSITFSFMSLCLVVG 780
Db 736 SLVGVPISSLQNLQPLAVDDQTYFTASLPFKNCGADHICDDLSVWFGFPDLKTLVVG 795
QY 781 GRPREFNVTVNRNGBDSRTQVTFPPDLDSYRKVSTLQ-----NQRSQSRW 829
Db 796 SDELELVDTVNSDGBSDSGTITVTFYPVGLSFRVVAEQVFLRKEDQOQRGQSHLH 855
QY 830 LACESASSTEVSGALKSTCSNHPITFPENSEVTNITFDVDSKASLGNKLLKANVTSE 889
Db 856 LMCD--STPDRSGLMSTSCSRHVIFRGSGQMTFLVTFDVPKABGLDRLLRLARVGE 913
QY 890 NNPRTNKTETFOLELPVKYAVMYVTSHGVSTKYLNFTASE-NTSRVMQHOYQVSNLGR 948
Db 914 NNVPGTPTKTFOLELPVKYAVMYTMSHDOFTKYLNFTSEKETSVEHREQVNNLGR 973
QY 949 SUPISLVFLVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNC 1008
Db 974 DVPVSINFVPIELKGEAVW-TMVSHQPQNLPTQCYRNLKPTQFDLLTHMQKSPVLD 1032
QY 1009 IAVCQRIQCDIPFFGQIEFNATLKNLSFDWYIKTSHNHLIVSTAELFNDVSFTLLP 1068
Db 1033 IADCLHLRCDIPSLGLDDELFTLKNLSFGWISQTLQKVLLSABEITFTSVYSQPL 1092
QY 1069 GOGAFVRSOTETKVEPPEPNPLPLTVGSSVGLLLALITAALYKLGFFKQYKDMSE 1128
Db 1093 QGEAFRAQTKTVLEMYKVNPPVPLVGVSSVGLLLALITAILYKAGFFKQYKEMLEE 1152

RESULT 6
QOYE7
ID QOYE7 PRELIMINARY; PRT; 1161 AA.
AC QOYE7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin.
SQ
SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.0%; Score 3236.5; DB 11; Length 1161;
Best Local Similarity 57.5%; Pred. No. 4.4e-232;
Matches 648; Conservative 163; Mismatches 303; Indels 13; Gaps 9;

QY 2 NLDTENAMTFOENARGFGQSVVQLGSRVWVGAPOEIVAAANQRGSLYOCDSYTGSCPIR 61
Db 21 NLDVEEPIVFREDAASFGQTVVQFGSRLVVGAPLEAVAVNQTGRLYOCAPATGQCPIV 80
QY 62 LQVPVEAVNMSLGLSLAATSPQLLACGPTVHTQCSNTYVYKGLCFLFGSNLRQOPKF 121
Db 81 LRSPLVEAVNMSLGLSLVATNNAQLLACGPTAQRACVKNMYAKGSCLLLGSSL-QFIQAV 139
QY 122 PEALRGCEQEDSDIAFLIDGSGSIIPHPFRKKEHWSTVMEQLKSKTLFSLMOYSEBFR 181
Db 140 PASMPECPQRQEMDIAFLIDGSGSINQRDFAQKDFVKALMGEFASTSTLFLMOYSNLIK 199
QY 182 THFTPEKQNNPNRSLIKPITQLGRTHATGLRKVKVRELFNITNGARKNAFKLILFLT 241
Db 200 THFTPEKNIIDPQSLVDPIVQLQCLTYTATGIRTVMEELFHSKNGSRKSKAKILLVIT 259
QY 242 DGEKFGDPLGYEDVTELDREGVIRYVIGVDAPRSEKSRQELNTVASKPRDRHVQINN 301
Db 260 DQKYRDPLEYSVDVIPAADKAGIIRYAGVDAFQEPALKELTNTIGSAPPQDHVKVGN 319
QY 302 FEALKTIONQRLKEKIFALEGTQSSSFEHMSQEGFSAATISNGPLLSVTGSDWAGG 361
Db 320 FAALRSIQRLQEKIFAIEGTQSSSSFEHMSQEGFSSALTSQDPLGAVGFSWSGG 379
QY 362 VFLYTSKESKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 421
Db 380 AFLYPNTRPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIFTQ 439
QY 422 NTGMWESNANVKGTOIGAYFGASCLSDVDVDSNGSTDVILIGAPHYYEOTRGQVSVCLP 481
Db 440 EARHWRPKSEVRGTQIGSYFGASCLSDVDVDRDGLVLI GAPHYYEOTRGQVSVFPVP 499
QY 482 RGQARWOCDAVLYGEOQOPWGRFCAALTLDVGNVGDKLTDVAIGAPCEEDNRGAVYLF 541
Db 500 -GVGRWQCEATLHGEQHPWGRFCAALTLDVGNVGDNLADVAIGAPCEESRGAVYIFH 558
QY 542 GTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 601
Db 559 GASRLIEIMPSPQRVTGSQLSURLQYFGQSLSGGQDLTQDGLVDLAVGQGHVLLRSUP 618
QY 602 VLRVKAIMEFPNPREVARNVFECNDQVVGKAGEVRVCLHVQKSTRDLREGQIQSVVTY 661
Db 619 LLKVELSIRFAPMEVAKAVYQCWERTPTVLEAGETVCLTVHKGSPDLL--GNVQSVRY 676
QY 662 DLALDSGRPHSRVNETKNSRRROTQVLGLTQTCETIKLQLPNCIEBPPVSVILRLNFS 721
Db 677 DLALDPGRLLISRAIFDETCKNTCTLRKTLGLGDHCHETVKLLPDCQVEDAVSPILLNFS 736

```

QY 722 LVGTPLSAFNGRLPVLAEQRLFTALFPPEKNCNDNICODDLSTFFSMLDCLVWVG 781
 DB 737 LVRDSASP-RNLHPVLAVGQDHTASLPEKNCCKQELLCGDLGFSNPSGLQVLVWVG 795
 QY 782 PRENVTVTVRNNGEDSYRTQVTFPPFLDLSYKVKVSTLQNRQSRWMLACESASSTEV 841
 DB 796 SPELTVTVTVMNEGEDSYGLVAFYYPAGLSYRVVTGTQ-QPHQYPLRLACEAPAAQED 854
 QY 842 GALKSTCSINHPIFFPENSVTENITFDVDSKSLGNKLLKANVTSENNMPTNKTEFQ 901
 DB 855 --LRSSCSINHPIFREGAKTTFMITFDVSYKAPLGRLLLRKASSENKPDNTKTAFF 912
 QY 902 LELPKYAVVMVTVSHGVSTKYLNFTASENTSR-VMQHOYQVSNLGSRLSPLISLVLVPV 960
 DB 913 LELPKYVTVTLISROQDSTNHVNFSSHGGRQEAHRVYVNNLSPLKLAVRVNFWVPV 972
 QY 961 RLNQTIVDRPQVTFSENLS--TCHTKERLPSHSDFLAELRKAPVNVNCSIAVQRCQDI 1019
 DB 973 LLNGVAVMD--VTLSSPAQGVSCVSKMPQPNPDFLTQIQRSSVLDSCSIADCLHFRCDI 1029
 QY 1020 PFGIOEFNATLKGNSLFWYIKTSHNHLIVSTAEILFNDVFTLLPGQGFVRSOTE 1079
 DB 1030 PSUDIODELFILRGNLSFGVSGTLOEKVLLVSEAEITFDTSVYSLPQGEAPLRAQVE 1089
 QY 1080 TKVEPPEVNPPLPLIVGSSVGGLLLLALITALAALYKLGFFKQYKQMM 1126
 DB 1090 TTLEEVYVPEPFLVAGSSVGGLLLLALITVWLYKLGFFKQYKEML 1136

RESULT 7

Q9WTV4 ID Q9WTV4 PRELIMINARY; PRT; 1161 AA.
 AC Q9WTV4;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Integrin alpha L.
 GN ITGAL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2J; TISSUE=Spleen;
 RA Ma R.2., Teuscher C.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF065902; AAD25885.1;
 DR HSSP; P20701; 1LFA.
 DR MGD; MGI:96606; Itgal.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 5.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VVFA; 1.
 SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209B431 CRC64;

Query Match 25.9%; Score 1523.5; DB 11; Length 1161;
 Best Local Similarity 33.9%; Pred. No. 2.9e-104;
 Matches 396; Conservative 220; Mismatches 461; Indels 91; Gaps 37;
 QY 1 FNLDENAMTFOENA-RGFQSVVQLOGSRVWVCAPOEIVAANRGSLYOCDYSTGSCPE 59
 DB 24 YNLDRTPQSLAQAGRHFGVQLQIEDG-VVVGAPGE---GDMTGLYHCRTSSEFCQP 79
 QY 60 IRLQVPVEAVNMSLGLSLATTPSPQLLACGPTVHQTCSNTYVYKGLCFLGSLNLRQOPQ 119

DB 80 VSLH--GSNHTSKYLGMTLATDAAGKSLACDPGLSRTCDQNTYLSGLCYLFPQSLGPM 138
 QY 120 KPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTKLSLMOYSSE 179
 DB 139 QNRPAVQECMKGVLDVLFIDGSQLDKDFKILEFMDKVMRKLNSNTSYQFAAQVSTD 198
 QY 180 FRIHFTFKEF-QNNPNRSLIKPIITQLLGRTHATGLRKVVRELFINTGARKAPKILF 238
 DB 199 CRTEFTFLDYVQKNKPNVLLGSVQPMFLTTTPRAINYVAHVPEESGARPDATKVLV 258
 QY 239 LITDEKFP--GDPYGEDVPELDEGVIRYVIGVDAPRSEKSKOELNLTASKPRDHV 296
 DB 259 IITDGEASDKGNISAAHD-----ITRYIIGIKGFVSKOKTLHIFASEPVEEFP 309
 QY 297 FOINNFEAKTTONOLREKIFAIEGTOTGSSSSFEHEMSOEGFSAITNSGPLLSTVGSY 356
 DB 310 KILDTFEKLDLFTDLORRIYAIETNRQDLTSFNMELSSSGISADLSKGHAVGVAGAK 369
 QY 357 DWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGA---IILNRNVQSLVLGAPRYQHTG 414
 DB 370 DWAGGFLLDLRELOGATFVGQPLTSDVRGGYLVTVAMTSSRSRPLLAAGAPRYOHVG 429
 QY 415 LVAMFR--QNTGWESNANVKGTQIAYFGASLCSVDVDSNGSTDVLVLICAPHYEQTRG 472
 DB 430 QVLLFOAPEAGRWNTQKIEGTQISYFGBELCSVDLDGGEAEILLIGAPLFGEQRG 489
 QY 473 GOVSVCPLPRGORARWQCDVLYGEOGOWPGRFGAALTVLGVDVNGDKLTDVAIGAPGED 532
 DB 490 GRVFTY---QRRQSIFEMVSELOQDGYPLGRFGAAITALTIDINGRLTDAVAGPLEB- 545
 QY 533 NRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLRYQFQSGLSGGQDITMDGLVDLTVGAQG 592
 DB 546 -QGVYIFNGKPG-GLSPQPSQRIQGAQVFGIRWFRGSIHGVDLGGDLADVVVGPBG 603
 QY 593 HVLLLASQPLRVLKALMEFNPREVARNVFECDQVQVKGKAG-EVRVCLVHVKSTDRLR 651
 DB 604 RVVVLSSREVVDVVTLSFSPEEIPVHEVECSYSAEEQKGVKLCFAIRKPLTPQ--F 661
 QY 652 EGOIQSVTVTYDLALDSGRHSRAVFNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPV 711
 DB 662 QORLLANLSYTLQDGHNRKSLPDPGSHLSNGTSITP-DKSCLDLFFHPICIQDLI 720
 QY 712 SPIVLRNLFSV---GTPLSAFNG-LRPVLAEDAQRLLFTALFPPEKNCNDNICODDL 767
 DB 721 SPINVSINFLLEEECTPRDQGRAMQPIRLPSIHTV-TKEIPFEKNCGEDKKCEANLTL 779
 QY 768 TFSFMSLCLVGGP-----REFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQN 821
 DB 780 SSPARS-----GPLRLMSSASLAVENTLSNGEDAYVWRDLDFPRLSFRKVMELQ- 831
 QY 822 QRSQSRWLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKL 880
 DB 832 --PHSRMPVSCBEL--TEGSSLLTKLKNVSSPIFKAGEVSLQVWFNTLLNSSWEDFV 887
 QY 881 LLKANVTSEN-NMPRTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHQ 939
 DB 888 ELNGTVHCENENSSLOEDNSAATHIPVLYPVNILLTKEQENSTLYISFTPKGPKTQQQV 947
 QY 940 YQV---SNLQORSLPISLVLVPLVRLNQTVIHDRPO--VTFSENLS-----TCHTKE-RL 989
 DB 948 YQVRIQPSAYDHNMP-TLEALVGP-----WPHSEDPITYTMSVQTDPLVTCHSDLKR 1000
 QY 990 PSHSDFLAELRKAPVNVNCSIAVCQRIQCDIPFGIOEFNATLKGNSLFDWYIKTSHNHL 1049
 DB 1001 PSSE---AQOCPUGV-----QPRCPVIF---RREILQVTGIVELSKELKAS-STL 1045
 QY 1050 LIVSTAEILFNDVFTLLPGQGFVRSQETTKVEPFEVNPPLPLIVGSSVGGLLALIT 1109
 DB 1046 SLCCSLSVSFNSKHPHLYGSKA-SEAQVLVAVKVDLIHEKMLHVYVLSGIGGLVLLFLIF 1104
 QY 1110 AALYKLGFFKQYKQYKMM-SEGPPGAP 1136

Db 118 SYKTEPDSYVRKDDALLKHVKHMLLTNTFGAINVYVATEVFRRELGARDPATKVLII 177
QY 239 LLTDGKFGDPLGVEDVPELDEGVIYVIGVDFAFRSEKSOELNTVASKPRDRHFQ 298
Db 178 IITDGE--ATDSNIDAADK-----IIRYIIGIKHFOFKESQETLHKFASKPASFFVKI 230
QY 299 INNFEALKTIONLREKIFAIBGTQTGSSSPHEMSQEGFAAITSNGPLLSSTVGSYDW 358
Db 231 LDTFEKLKOLFTELOKKIYIEGTSKQDLTSFNMELSSSGISADLSRGAHVAVGAKDW 290
QY 359 AGGVF-LYTSKESKSTFINMTVDSDMNDAYLGAAA-IILNRVQSILVGAPEYOHIGLV 416
Db 291 AGFLDLKADLQDDTFIGNELPTEVRAGLVGVTVTWLPSRQKTSLLASGAPRYQMGVRV 350
QY 417 AMFR--QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQ 474
Db 351 LRFQEPGGGHWGSOVQTIHGTOIGSYFGELCGVDVDQDGETELLIGAPLFFGEORGGR 410
QY 475 VSCPLPRQGRARQCDV--LYGEOGQPMGRGAALTVLGDVNGDKLTDVAIGARBEED 532
Db 411 VFYI-----QRRQLGFEVSELOQDPCYPLGRFCEAITALTIDINGDGLVDVAVGAPLEE- 464
QY 533 NRCAYVLFHGTSGSGISPSHSORAGSKLSPRLOYEQSLSGQDLMGLVDLTVGAQG 592
Db 465 -QCAVIFNGRHG-GUSPOPSQRIEGTVLSGIGQWFGRSIHGVKDLGGDLADVAVGAS 522
QY 593 HVLLLRASQPLVRKVAIMEFNPREVARNVFECNDQVV-KGKEAGEVRCVLHVQKSTRDLR 651
Db 523 QMIVLSRPPVDMVTLMSFPAIPVHEVECSYSTSKMKKEGVNIICQI-KSLIPQF- 580
QY 652 EGOIQSVVYVDLALDGRPHSRVAFNETKNSTRQTVGLGTQTCETLKLQLPNCIEDPV 711
Db 581 QGRVANLTYTLQDGHRRRRGLFPGRRHRLRNIAVT-TSMSCDTDFSPHFVPCVDLI 639
QY 712 SPVLRLNLSL-----VCTPLSAFGN-----LRPLAEDAQRLLFTALFPPEKNCNDNICQ 762
Db 640 SPINVSLSLWEEEGTPDRQAGKOIPLRLPSLHSETWEI-----PFEKNGEDKKCE 694
QY 763 DLSLITFSMSLDCLVGGPREFNVTVVRNDEGDSYRQTVTFPPFLDLSYRKVSTLQW 822
Db 695 ANLRVSFSPARSALALTAFAISLVELSLNLEEDAYWQLDLHPFGGLSFRKVEML--- 751
QY 823 RSORSRLACES--ASSTVSGALKSTSCSINHPIPPENSEVTNTTFDVSCKASLGNKL 880
Db 752 KPHSQIPVSCPEELPEESRLISRAL--SCNVSSPIFKAGHSVALQWMENTLVNWSGDSV 808
QY 881 LKANYTSENN---MPTNKTEFQLELPVKYAVVWVTSHGVSSTKYLNFTASENTRVM 936
Db 809 ELHANYTCNNEDSULLEDNSATTI---IPILYPINILIQOEDSTLYVSFTPKGPKHQV 865
QY 937 QHOYQV--SNLQORSIP--ISLVFLVPLVRLNQTVIWRDPQVTFSENLSSTCTK--ERLP 990
Db 866 KMYQVRIQBSIHDNIPTLEAVGVQPPSEGPITHQWSVQMEPPV--PCHYEDLERLP 923
QY 991 SHSD--FLAELRKAPVNVCSIAVCORIQCDIPFFGIQEBFNATLKNLSFDWYIKTSHNH 1048
Db 924 DAAEPCLPGALPCPVV-----FRQELVQVIGTLELVGGEIAS-SM 964
QY 1049 LLIVSTAELFNDVSFTLLPGCAFVRSQTETKVPEFEPVNPPLIVGSGVGLLLALLI 1108
Db 965 FSLCSSLSISFNSKHFHLYGSNASL-AQVVMKVDDVYKQMLLYLVLSVIGIGLGLLLLI 1023
QY 1109 TAALYKLGFPKQYKDMWSEG-GPPGAEP 1136
Db 1024 FIVLYKVGFKRNLKEMERAGVPPNGIP 1052

RESULT 11

Q98TF0

ID Q98TF0 PRELIMINARY; PRT; 1187 AA.

AC Q98TF0;

DT 01-JUN-2001 (TReMBLrel. 17, Created)

DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE CD11-2.
GN CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048537; BAB39135.1; .
DR HSSP; P20701; 1LPA.
DR InterPro; IPR001969; Asprotease site.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;

Query Match 22.9%; Score 1344.5; DB 13; Length 1187;

Best Local Similarity 31.3%; Pred. No. 7e-91;

Matches 373; Conservative 209; Mismatches 453; Indels 157; Gaps 43;

QY 1 FNLDTENAMTFQENARG-FCQSVVQLQ-GSR--VVVGAPQEIIVANQORSLVQCDYSTGS 56

Db 32 FNIDTEHPLURFNGTPEDFEGSYVQTEFGNRKQIIVGAPLE---GNSAGEMYSCCTADLOS 88

QY 57 CSPIRLQVP---VEAVNMSLGLSLAATSPQLLACGPTVHTQTSNTYVVKGLCLFLGS 112

Db 89 CK-RLQRPQSGSVRFFGNSAAVSSAALTS-----CSPYFAHECDGNSYLVNGVCYQFNS 140

QY 113 NLRQOPQKPEALRQCPQSDSIAFLIDGSGSIIIPHFRMKEWSTVMEQLKSKTLFS 172

Db 141 SL-QAVSNFTAAAYQECSEKREVLNLFVFDGSSMKAVEFDMKNFKIDVMKKLSNSSIKFA 199

QY 173 LMQYSEEFRIHFTTFEFQNNPNRSLIKPITOLLGRTHATGLRKVVRELFN-ITNGARK 231

Db 200 AVQFSTEIRTVDFPDYQNGSAEELMKE-RHMKSLTNTYKAIYNYLVKNLVNSVSSGADP 258

QY 232 NAFKILFLLLTDCGKFGDPLGYED--VIPELDRGVIRYVIGVDAPFRSEKSRQELNVTAS 289

Db 259 NAQKALVIITD---GDPSDDNDYINILCDQNILRYIIGVG-----KVDLTUTQLAA 309

QY 290 KPPRDHVFOINNFEALKTIQNLREKIFAIBGTQTGSSSSFEHMSQEGFAAITSNGPL 349

Db 310 EPKLNNTFYIQEYNGKLGILLDNLQKKIYNIEGSKEAHGRDQKELSQSGFSVVYQESVI 369

QY 350 LSTVGSYDWAGVFLYT---SKEKSTFINMTVDSMDNDAYLGAAAAIILNRVQSLVLG 406

Db 370 VGSVGSNDWRGALYEVMSGSKFRQTEITDPAVN---KDSYMGYSTVLGMRHGVSLLFSG 426

QY 407 APRYOHIGLVAMFRONTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDVLVIGAPHY 466

Db 427 APRAEHTLVLTLTKNSTWTVNRNINGEQISYFGASLUSLDDVSDGSDFLLVGAPLP 486

QY 467 YE-QTRG-QQSVCPPLPRQARWQCDVLYGEQ--GQPWGRFGAALTVLGDVNGDKLTD 522

Db 487 YQSQPRAEGRLYVYTL-----SEQYSQKTLQSTTGRFATSLASLKDNLGDLSD 535

QY 523 VAIGAPGEEDNRGAVLPHGTSGSGISPSHS--QRTAGSKLSRQLYFGQSLSGQDLTMD 581

Db 536 VAVGAPLE--NEGVMYIYLDGTHGINPEHAPQIPARSVLPGLQGFVSLSGQDMND 593

Qy	582	GLVDLTGVAQGHVLLRSQVLRVKAIIMEFNPREVARNVFECNQDVYKKGKEAGEVRVCLH	641
Dd	594	NLPDVIIVTGGIIVLLNARVMVSQAQLSPNMEISLNYEPCGS--NAPNAFNALTSCTF	651
Qy	642	VQKSTRDLREGIOISV--VTYDIALDSGRPHSRVAFNETKNSTR--RQTQVLGLTQTCE	697
Dd	652	VTERTS--TGSLEKKLVLSNLNDLVWVGMSRGFFDPMDSSRTLQQOSVLLDSSSCS	708
Qy	698	TLKLQPLNCIEDPVSPILRLNFS---LVGTPLSAFGNLRPVLADDAORLFALTLPFEK	753
Dd	709	NFSIFMLRCVADIVSPCLKIRWNFSOTQLMSGNSLAFL-DIQSTEEVEVFL-----	761
Qy	754	NCGNNDNICODDLGITFTSFMSLDCLLVGGPREFNVTVTRNGEDSYRTQVTFPPFDLSY	813
Dd	762	NC-NSNSCVADLKLPFSFTN-DTLVVENQAHTVLVSLANPGDDSYNTSVILHYPEGLSL	819
Qy	814	RKVSTLQNQRSQRWRACBSASSTEVSFGALKTSCSINHPIPFENSEVTFNITFDV---	870
Dd	820	SKFDAIKPSRTR-----SSCDDRDSGATNRITCSINLPVYRSRGTTOFLGTFRVTKW	871
Qy	871	DSKASLGKLLKANVTSENNMPRTNKTEFOLELPVKYAYVMVVTSHGV-STKYLNFTAS	929
Dd	872	DYDWSDRMETITANSDNNGM---SDMSVRRSIPQVFAVELAISLVAEDSVTYLNFSL	928
Qy	930	ENTSRVMOHQVOYSNLGORSLPISLVLVPVRLNQTVIMDRPQVTFSENLSSTCHKERL	989
Dd	929	DRGPKPLNIYKVNGLGKLPVSVTLSPCQ-----TTHVTLTPHNFSMQ	974
Qy	990	PSHSDFLAELRKAPVNC-----SIAVCRIQCIDIFFGLQIE	1027
Dd	975	EVBHSFTSSYHQ--IIMCLINKHLFPSPELSAVQTRTTGRSLWC-----VSSI	1026
Qy	1028	FNATLKGNL-----SFDWKVKTSHNHLLIVSTAETILFNDSVFTLLPG-QGA	1072
Dd	1027	FRSSV--NLMAEAVLQNVKEYESKYSFY-EFRRDHVFNIS-AELNFNTSYNSTGLKYN	1082
Qy	1073	FVRSQTEKVEPFVPPNPPLPIVSGSSVGGLLILLALIITAALYKLGFFKRYOKD	1124
Dd	1083	PHRSQTEKVEFVIPPBRMLIVCTAGVGGFLLIIILLLLCKGFFKENRPD	1134
 RESULT 12			
Q8HZVO	ID	Q8HZVO	PRELIMINARY; PRT; 927 AA.
AC	Q8HZVO;		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	Lymphocyte function-associated antigen 1 (Fragment).		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Bos.		
OX	NCBI TaxID=9913;		
RN	{1}		
RP	SEQUENCE FROM N.A.		
RA	Thumbikat P., Kannan M.S., Maheswaran S.K.;		
RT	"Sequence of the alpha subunit of bovine lymphocyte function-		
RT	associated antigen 1."		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBAJ databases.		
DR	EMBL; AF440778; AAK63636.1; -		
NON TER	1		
FT	NON TER	927	927
SQ	SEQUENCE	927 AA; 102523 MW;	02E3CF09917081EC CRC64;

Qy	234	FKILFLLTDEKPGDPLGVEDVIPLELDREGVIYVLVGDAFRSEKSROELINTVASKPPR	293
Dq	64	TKVLIITIDGK---PPTNTLMRPKTSS----RSLLGIGNFKTKESQBALHQFAKPYVE	116
Qy	294	DHYFOJNNFEALKTIQNOLREKIFAJETGTOTGSSSFEHEMSQEGFSAAITSNGPLLSTV	353
Dq	117	EFVKIIDLTEFKLKDLETELQKKLYIECTSKQDLTSFNMELESSGISADLSSEHGUVGVAG	176
Qy	354	GSYDWAGGVF-LYTSKEKSTFFINMTVRDSMDNDAYLGYAAA-IILNRNVOSLVLGAPRYO	411
Dq	177	GAKMWAGGFLLDKADLKSTTFVUGNEPLTVESRAGYLGYTVTRLPSPRGTMSSLATGAPKYQ	236
Qy	412	HIGLVAMFRQ--NTHGMWESNAVUKTOIGAYFCASICSVDVDSNGSTDVLVLGAPHYEQ	469
Dq	237	HVGRVILLFOQPKRGPMWSQIQEIDGQIGSYFGBELCGVDVDRDGSETLELLIAAPLYYE	296
Qy	470	TRGOVSVCPLPRGARWCODAVLYGEQOPMGRFGAALTVLGDVNGDKLTDOVAIGAPG	529
Dq	297	QRGRVFYI---OKIQLEFOMVSELOGETGYPLGRFGAIAALTINDGBELTDVAVGAPL	353
Qy	530	EEDNRGAVLYPHGTSGSGISPSHSQRSAGSKSPRIQYFGQSLSGGODLTMOGLVDLTVG	589
Dq	354	EE--QGAVYIFNGQQ-GLSPRESQRIEGTQMPSGIQWFGRSIHGVKDLGGDLADVAVG	410
Qy	590	AQHVVLLRSQPVLRVKAIMEFPNPREVARNVPFCNDQOVKGKAG-EVRVCLHVQK--ST	646
Dq	411	AEQGVIVLSSRPVVDIITSVSFPAIPVHEVECSYSTSNQKKEGNVLTVCFOVKLSUIT	470
Qy	647	RDLREGIOGSVVTYDLALDSGRPHGRAVENETKNSTRROTQVLGLTOTCETLKLQLPNC	706
Dq	471	-----FQGHVANLITYTLQDCHTRERGLEPFGKHKLIGNATVPV-KSCFVFWFHFPIC	525
Qy	707	IEDPVSPVILRNFLS---VGTPLS--AFGNLRPVLAEDAORLFTALFPEKCNKGNDNIC	761
Dq	526	IQDLJSPIVNSLSYWELEGTPROPRALORDIPPILKPSPLETKEIPEKNCGEDKNC	585
Qy	762	QDLSITFSFMSLDCLVGGPRENVTVTRNGEDSRYQTQVTFPPPLDLSYRKVSTLQN	821
Dq	586	EADLKAFSDMRSKILRLTFSASLVRLTLURNAEDAYWQVTLSPFQGLSRFKVEL--	643
Qy	822	QRSQRWLACESASSSTEVSALKSTSCSINHPIFPENSEBVTENITFDVDSKASLGNKLL	881
Dq	644	-KPHSHVPVCEELPEAVVHS-RALSINVSSIIFGEDSMWDIQVMENTLOKSWGDFIE	701
Qy	882	LKANVTS-----ENNMPRTNKTEFQLEPLPKYAVYVVTVSHGVSTKYLNFTASENTS	933
Dq	702	LQANVSCNNEDSILLEDNSATTS-----IPVMYPINVLTKQENSTLYISFTPKSPRI	754
Qy	934	RVMOHQVOV-----SNLCORSPLTSLVFLVPVRLNOTVI---WD---RPQVTF-S	982
Dq	755	HVHXHIYQVRIQPSNYONMP-PLEALVRVPVRHSEGILTHKWSIQMBPPVNCSPRNLES	813
Qy	983	CHTKERLPSSHDFLAELRKAPVNVCSIAVCORIQCIDIFFGIGOEFNATLKGNSLFDWI	1042
Dq	814	SDEAE-----SCSGT--EFRCPIDF---ROEILVQVNGMVELRGTI	850
Qy	1043	KTSNNHLLIYSTABILENDSVFTLLPQOGAFVRSQETKVEPPEVPNPLFIIVGSSVGL	1102
Dq	851	KAS-SMLSSCLSLAISNSSKHFFHLHGRNASM-AQVVMKVLDLYVEKMVLYLYLSGIGGL	908
Qy	1103	LLLLIIITAALYKLGFFKR	1120
Dq	909	LLLLLIFIATALYKVGFFKR	926

RESULT 13	
OB8340	
ID	OB8340
AC	OB8340;
DT	01-NOV-1998 (TRENBLrel. 08, Created)
DT	01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brennan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020045; AAC23662.1; -
DR HSSP; P11215; IBHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWF; 1.
SQ SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match 19.6%; Score 1154.5; DB 11; Length 1167;
Best Local Similarity 28.9%; Pred. No. 1e-76;
Matches 356; Conservative 214; Mismatches 458; Indels 205; Gaps 43;

QY 1 FNLDTEA--MTFQENARGGQSVVQLQGSVVVGAPQEIIVAAVNRGSS-----LYQCDDY 52
DB 20 FNNVDWAWYALQCAPAVLSLLHLDPSN-----NOTCLLVARRSSNNTAALYRCAL 74
QY 53 STGSCPIRLQVPEAVNMSLGLSLAATT--SPQLLAC-GPTVHOTCSNTVYKGLCLF 109
DB 75 SI-SPEIACQ-PVEHICMEKRGYQGVTLVGNHNGVLVCIVQVQARFRSLNSELTCACSL 132
QY 110 FGSNLQAPQKPFELARG-----C-----PQE 131
DB 133 LTPNLQQAQYPSDLGFLDIPGAHVDSGYCRSKGSGTGEEKSARRRTVEEED 192
QY 132 DSDFIADLGGSGIIPHDFRRMKEMWSTVMEQL--KSKTLFSLMOYSEEFRIHFTKEF 189
DB 193 GTETAIVLDSGSIERSDFQAKNFISTMWRNEYEKCFECNFALVOYGAIVQTEFDLQES 252
QY 190 QNNPNRSLIKPTQLLGRTHRTATGLRKVVRELFNITNGARKNAFKILFLLTGDGKFGDP 249
DB 253 RDINASLAKVQSVQVKEVTKTASAMQHVLDNIFIPSRGSRKALKVMVVLTDGDI FGDP 312
QY 250 LGYEDVIPELDRGVIRYVIGVGDAPRSEKSRQELNATVASKPPRHVFOINNFEALKTQ 309
DB 313 LNUITVINSFKMGQVRFAGVGDGAFKNNNTYBELKLIASDPKEATFTKVTNYSLDGLL 372
QY 310 NQUREKIFATEGTQSGSSFEHMSQEGFSAAITSNGP--LLSTVGSYDWAGVFLY-TS 367
DB 373 SKLQQRIVHMEGT---VGDALQVLAQTGSAQLDKGVLLTGVAFNWSGGALLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DMNDAYLGAAATILNRVQSLVGLAPRYOHIGLVAMFRQWTGM 425
DB 430 NGRGCFNLQAKEDSRVTQVSYLGSVLAVLHKAHGISYVAGAPRHKLKRGAVPELRKEDR- 488
QY 426 WESNA---NVKGTQICAYFGASLCSDVDNSNGSTDLVLGAPHYEQTGGQGVSCPLPR 482
DB 489 -EDAFVRRIEGQMGSYFSGVLCVPDIDMDGTTDFLLAAAPYHIRGEGRVYVQVPE 547
QY 483 GQARWQCDAVLGEGQOPNGRFGCAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSGHFGLTNSRFGFMAAVGDIQDKFTDVAIGAPLEGFAGDGASYS 606
QY 537 VYLFHGTSGSGISPSHSQRIAGSKPLRQYFQGSLSGGQDLTMDGLDVLTVGAQGHVLL 596

DB 607 VVIYNHSG-GLYDPSQOIRASSVGLHYFGMSVSGGLDFNGGLADITVGSRDSAVV 665
QY 597 LRSQPVLRVKAIMEFNPVARNVFECDNVVYKGEAGEVRVCLHVQKS---TRDLREG 653
DB 666 LRSRPVVDLTVSMTEFP-----DALPMVFIGM--DVNLCFEVDSSVWASEPGLREM 715
QY 654 QIOSVVTYDLDALDSGRPHSRVFNKTSRTRQTVLGLTQTC----- 696
DB 716 FLNFTVDVDV-----TKQRRLQCEDSSGCCLRWKNGSGFLCEHFWLI 760
QY 697 ETLKQLPNCIEDPVSPIVLRNLSVGLTPTGFLSAFNLN-----PVLAEQAQLFTALP--P 750
DB 761 STEEL-----CEEDCFSNITIKVTYE-----FOTSGRRDYPNPTL--DHYKEPSAIFQLP 809
QY 751 FEKNGCNDNIQDDLSITFSFMSLDCLVVGGPREFNVTVTVRNDGEDSVYRTOVTFEPLD 810
DB 810 YEKDCNKVFCIAEIQLTN-ISQELVVGVTKEVTMNISLTNSGEDSVYRTOVTFEPLD 868
QY 811 LSYRKVSTLQONORSQSRWRLACESASSTEVSALKSTSCSINHPIPPENSEVTNITFDV 870
DB 869 LQFKKI-----QKPVSPDVQCDPKPV---ASVLVMNCKIGHPII-KRSSVNVSVTMQL 918
QY 871 DSKASLGNKLLKANVTSENMPRTNKTEFQLELPVKYAVYVMTSHGVSTKYLNTASE 930
DB 919 EESVFPNRTADITVTISNSNEKSLARETR---SLOFRHAFIAVLRS--PSVMYMN--TSQ 971
QY 931 NTSRVVMOHQVQSVNLGQSLPISLVFLVPRLNQTVINDRPOVTFSENLSST-----CHT 985
DB 972 SFSDHKEFPFNHNGENLFGAVFQLOICVPIKLODF-----QIVRVKNTKTQDHTECTQ 1025
QY 986 KERLPSSHDFLAELKAPVNVCSIAVCQRIQCDIPFGIQQEBFNATLKNLSFDMYIKTS 1045
DB 1026 SQEPACGSDPVQHVKEHWSVCAI-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLLIVSTA-----EILPNDVSVTLPLPQCAFVRSQTETKVEFP-----EVPNPLPLIV 1095
DB 1066 HTKQLLRDVELPILGEISFNKSLYEGLNAE-----NHRTKITVIFLKEEETRSLPLII 1119
QY 1096 GSSVGGLLLLALITAALYKLGFFKQYKDMNSE 1128
DB 1120 GSSIGGLLVLVIIAILFCGFFKRYQQLNLE 1152

RESULT 14
O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brennan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -
DR HSSP; P11215; IBHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.

GenCore version 5.1.6
Copyright (c) 1993 - 2003: Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.2978 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481b-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMWSEGGPGGAPFP 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries:

Database : A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	5852	99.5	1153 19	Human Beta-integrin
2	5852	99.5	1153 21	Human CD11b protei
3	5852	99.5	1153 23	Human Beta2 integr
4	5852	99.5	1153 23	Human integrin 1 a
5	5852	99.5	1153 23	Integrin Mac-1 alp
6	5842	99.3	1153 11	Alpha subunit of M
7	3450	58.6	1163 11	p150.95 alpha subu
8	3436	58.4	1163 19	Human Beta-integrin
9	3436	58.4	1163 21	Human CD11c protei

10	3436	58.4	1163	23	ABG61470	Human Beta2 integr
11	3434	58.4	1163	24	ABU07406	Protein differenti
12	3401	57.8	1161	18	AAW78166	Human beta-2 integ
13	3401	57.8	1161	18	AAW23049	Human beta-2 integ
14	3401	57.8	1161	19	AAW72825	Human alpha-d. Ho
15	3401	57.8	1161	19	AAW65089	Human Beta-integrin
16	3401	57.8	1161	19	AAW57491	Human Beta2 integr
17	3401	57.8	1161	20	AAW73342	Human alpha d clon
18	3401	57.8	1161	21	AAW07359	Human alpha d prote
19	3401	57.8	1161	23	ABG61468	Human Beta2 integr
20	3385.5	57.5	1161	18	AAW23064	Human beta-2 integ
21	3385.5	57.5	1161	19	AAW72837	Human alpha-d deri
22	3385.5	57.5	1161	19	AAW65106	Human Beta-integrin
23	3385.5	57.5	1161	20	AAW73343	Human alpha d prote
24	3385.5	57.5	1161	21	AAW07376	Human alpha d prote
25	3385.5	57.5	1161	23	ABG61485	Human Beta2 integr
26	3226.5	54.8	1161	16	AAW78169	Rat alpha-d subuni
27	3224.5	54.8	1161	18	AAW23062	Rat beta-2 integrin
28	3224.5	54.8	1161	19	AAW72824	Rat alpha-d #1. R
29	3224.5	54.8	1161	19	AAW60004	Rat alpha d polype
30	3224.5	54.8	1161	21	AAW07374	Rat alpha d protei
31	3224.5	54.8	1161	23	ABG61483	Rat Beta2 integrin
32	3217.5	54.7	1161	19	AAW65104	Rat beta-integrin
33	3217.5	54.7	1161	20	AAW73345	Rat alpha d integr
34	3212	54.6	1161	20	AAW23061	Rat alpha d #2
35	3212	54.6	1161	19	AAW72836	Mouse alpha-d #2
36	3212	54.6	1161	19	AAW65103	Mouse beta-integrin
37	3212	54.6	1161	19	AAW60003	Mouse alpha d poly
38	3212	54.6	1161	20	AAW73347	Mouse alpha d prote
39	3212	54.6	1161	21	AAW07373	Mouse alpha-d prot
40	3212	54.6	1161	21	ABG61482	Mouse Beta2 integr
41	3208	54.5	1161	16	AAW78168	Mouse alpha-d subu
42	3203.5	54.4	1151	18	AAW23059	Rat beta-2 integrin
43	3203.5	54.4	1151	19	AAW72834	Rat alpha-d #2. R
44	3203.5	54.4	1151	19	AAW65101	Rat beta-integrin
45	3203.5	54.4	1151	19	AAW60001	Rat alpha d polype

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; Protein; 1153 AA.

XX
AC AAW65090;

XX
DT 28-SEP-1998 (first entry)

XX
DE Human Beta-integrin CD11b subunit protein.

XX
KW Beta-integrin alpha-d subunit; human; modulator; treatment: psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.

XX
OS Homo sapiens.

XX
PN US5728533-A.

XX
PD 17-MAR-1998.

XX
PF 07-JUN-1995; 95US-0485618.

XX
PR 07-JUN-1995; 95US-0485618.

XX
PR 23-DEC-1993; 93US-0173497.

XX
PR 05-AUG-1994; 94US-0286889.

XX
PR 21-DEC-1994; 94US-0362652.

XX
PA (ICOS-) ICOS CORP.

XX
PI Gallatin WM, Van DER VIEREN M;

DR WPI; 1998-206565/18.
XX Screening assay for modulators of integrin binding - using
PT immobilised or labelled alpha-d polypeptide, useful for, e.g.
PT treating type-I diabetes
XX
XX Example 5; Fig 1A-D; 106pp; English.
XX This sequence represents a human beta-integrin CD11b subunit which is
CC used to describe a method for identifying compounds that modulate the
CC interaction of the beta-integrin alpha-d subunit with a binding partner
CC of alpha-d which involves contacting an alpha-d polypeptide with an
CC alpha-d binding partner, one of which is immobilised and the other of
CC which is labelled, in the presence of a test compound, and determining if
CC the compound affects binding between the alpha-d polypeptide and alpha-d
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress
CC syndrome and rheumatoid arthritis.
XX
SQ Sequence 1153 AA;
Query Match 99.5%; Score 5852; DB 19; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY 1 FNLDENAMTQENARGGQSVQVQGGSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDENAMTQENARGGQSVQVQGGSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATSPQALLACGPTVHQTCSNTYVKGCLFLFGSLNRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPQALLACGPTVHQTCSNTYVKGCLFLFGSLNRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGIIIPDFRMKEVSTVMEQLKSKTFLSLMQYSEEP 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGIIIPDFRMKEVSTVMEQLKSKTFLSLMQYSEEP 196
QY 181 RIHFTFEKFNPNRSLKPIITQLLGRTHATGLRVVRELFTNGARKNAFKILFL 240
Db 197 RIHFTFEKFNPNRSLKPIITQLLGRTHATGLRVVRELFTNGARKNAFKILV 256
QY 241 TDGEKFGDPLGYEDVPELDEGVIRYVIGVDGAFRSEKSKQELNTVASKPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPELDEGVIRYVIGVDGAFRSEKSKQELNTVASKPRDHVFQIN 316
QY 301 NFEALKTIQNLREKIFAEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAG 376
QY 361 GVFLYTSKSKSTFNTMTVDSDMDNDAYLGXAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFNTMTVDSDMDNDAYLGXAAAIIILNRVQSLVGLGAPRYOHIGLVAMFR 436
QY 421 QNTGWESNANVKTQICAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 480
Db 437 QNTGWESNANVKTQICAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDVLYGEGQPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLOYGQSISGGQDLTMQGLVDLTGVAQGHVILLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLOYGQSISGGQDLTMQGLVDLTGVAQGHVILLRSQ 616
QY 601 PVLVRKATMEFNPREAVNPFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKATMEFNPREAVNPFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661: YDLALDSGRPHRAVFNETKNSRRTQVLTGTCTETLKLQLPNCIEDPVSIVLRINF 720

Db 677 YDLALDSGRPHRAVFNETKNSRRTQVLTGTCTETLKLQLPNCIEDPVSIVLRINF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPFENKNCNDNI CODDLSITTSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPFENKNCNDNI CODDLSITTSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLTQNRQSQRSLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLTQNRQSQRSLACESASSTEV 856
QY 841 SGALKSTSCSINHPIEPENSEVTNITFDVDSKASLGNKLLKANTYSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIEPENSEVTNITFDVDSKASLGNKLLKANTYSENMPRTNKTFF 916
QY 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTRVVMQHOVQVSNLQQRSLPLSLVLPV 960
Db 917 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTRVVMQHOVQVSNLQQRSLPLSLVLPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036
QY 1021 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGQAFVRSQTET 1080
Db 1037 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGQAFVRSQTET 1096
QY 1081 KVEPFEPVNPPLPIIVGSSVGLLLALITAAALYKLGFFKRYQKDMMSSEGPPGAEPQ 1137
Db 1097 KVEPFEPVNPPLPIIVGSSVGLLLALITAAALYKLGFFKRYQKDMMSSEGPPGAEPQ 1153
RESULT 2
AAB07360
ID AAB07360 standard; Protein; 1153 AA.
XX
AC AAB07360;
XX 17-JAN-2001 (first entry)
XX Human CD11b protein sequence.
XX Human; macrophage infiltration inhibition; alpha_d integrin;
XX leukocyte integrin; Leu-CAM; leukointegrin; immune response;
XX inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
XX atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
XX lung inflammation; acute respiratory distress syndrome; Crohn's disease;
XX rheumatoid arthritis; central nervous system injury; CD11b.
XX Homo sapiens.
XX WO200029446-A1.
XX 25-MAY-2000.
XX 16-NOV-1999; 99WO-US27139.
XX 16-NOV-1998; 98US-0193043.
XX 08-JUL-1999; 99US-0350259.
XX (ICOS-) ICOS CORP.
XX Gallatin MW, Van Der Vieren M;
XX WPI; 2000-387751/33.
XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
XX macrophage infiltration and reduce inflammation at central nervous
XX system injury sites -
XX Example 5; Fig 1; 270pp; English.
XX Integrins are a class of membrane-associated molecules that participate

in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha d. The alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type 1 diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (IAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

Query Match 99.5%; Score 5852; DB 21; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQGVVQLQGSRRVVGAPQEIIVAAHQRSGLYQCYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFQGVVQLQGSRRVVGAPQEIIVAAHQRSGLYQCYSTGSCBPI 76

QY 61 RLQVPVEAVNMSIGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQOPOK 120
DB 77 RLQVPVEAVNMSIGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQOPOK 136

QY 121 FPEALRGCPQEDSDIAFLDGSIIIPHDFFRMKEWVSTVMEQLKKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLDGSIIIPHDFFRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 196

QY 181 RHFTFKFQNNPNRSLIKPTOLLGRTHATGIRKVVRELFNITNGARKNAKILVVI 240
DB 197 RHFTFKFQNNPNRSLIKPTOLLGRTHATGIRKVVRELFNITNGARKNAKILVVI 256

QY 241 TQGEKFGDPLGVEDVPELDRGVIRYVIGVDADFSEKSRQELNTAVSKPRDHVQIN 300
DB 257 TQGEKFGDPLGVEDVPEADRGVIRYVIGVDADFSEKSRQELNTAVSKPRDHVQIN 316

QY 301 NFEALKTIQNLREKIFALEGQTQSSSFEHEMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFALEGQTQSSSFEHEMSQEGFSAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINNTRVDSNDAYLGYAAAILRNVRQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSNDAYLGYAAAILRNVRQSLVLGAPRYQHIGLVAMFR 436

QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICGAPHYEQTRGGQVVCPL 480
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICGAPHYEQTRGGQVVCPL 496

QY 481 PRQQRARMCQDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAICAPGEEDNRCAGVILF 540
DB 497 PRQQRARMCQDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAICAPGEEDNRCAGVILF 556

QY 541 HGTSSGSIQSPHSQRIAGSKLSPRLQYFQCSLSGGQDLTMDGLVDLTGVAQGHVLLLRSQ 600
DB 557 HGTSSGSIQSPHSQRIAGSKLSPRLQYFQCSLSGGQDLTMDGLVDLTGVAQGHVLLLRSQ 616

QY 601 PVLRLVKATMEFPREVARNVFCNDQVKGKEGVRVCLHVOKSTRDLREGQIOSVVT 660
DB 617 PVLRLVKATMEFPREVARNVFCNDQVKGKEGVRVCLHVOKSTRDLREGQIOSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGTQTCETLKLQPLNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSRROTQVLGTQTCETLKLQPLNCIEDPVPVILRLNF 736

QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICQDDLSITFSFMSLCLVVG 780

DB 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICQDDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRQVTFPPFDLSYKSVSTLQORSORSNRLACESASSTEV 840
DB 797 GPREFNVTVVRNDEGDSYRQVTFPPFDLSYKSVSTLQORSORSNRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916

QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASNTSRVMQHOYQVSNLQORSLSLVLVLPV 960
DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASNTSRVMQHOYQVSNLQORSLSLVLVLPV 976

QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVVCNCSIAVCORIQCDIP 1020
DB 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVVCNCSIAVCORIQCDIP 1036

QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSQTET 1080
DB 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSQTET 1096

QY 1081 KVEPPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
DB 1097 KVEPPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 3
ABG61469
ID ABG61469 standard; Protein; 1153 AA.
XX
AC ABG61469;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human Beta2 integrin alphaCD11b subunit.
XX
KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX
OS Homo sapiens.
XX
PN WO200230980-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US32059.
XX
PR 13-OCT-2000; 2000US-0688307.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin WM, Van Der Vieren M;
XX
DR WPI; 2002-463260/49.
XX
PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -
XX
PS Example 5; Page 191-194; 270pp; English.
XX
CC The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an

CC domain polypeptide comprising at least one disulfide bond so that the
CC domain is stabilised in a desired conformation. The polypeptide of
CC the invention may have anti-inflammatory or immunosuppressive activities.
CC The polypeptides of the invention have an open conformation and are
CC useful as immunogens to produce antibodies that selectively bind to
CC integrin I-domain; and for identifying a modulator of integrin activity,
CC or of interaction of an integrin and a cognate ligand. The polypeptide
CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
CC is useful for treating or preventing an integrin mediated disorder which
CC is an inflammatory or autoimmune disorder in a subject and for
CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
CC disease, nephritis, human immunodeficiency virus (HIV), myocardial
CC infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis.
CC A therapeutic composition comprising the peptide of the invention is
CC useful for treating an integrin mediated disorder in a subject. The
CC polypeptides and/or active or antigenic fragments are useful as
CC reagents for diagnosis of integrin-mediated disorders. The present
CC sequence represents the human integrin-1 alpha-M protein subunit used to
CC generate the mutant polypeptides of the invention.

XX SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 23; Length 1153;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEI VAAHQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEI VAAHQSGSLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTQCSNTYVKGCLFLFGSLNRQQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTQCSNTYVKGCLFLFGSLNRQQPOK 136
QY 121 PPEARLGCPOEDSDIAFLDGSGLIIPHDPRMKWVSTVMEQLKSKTFLSLMOYSBEF 180
DB 137 PPEARLGCPOEDSDIAFLDGSGLIIPHDPRMKWVSTVMEQLKSKTFLSLMOYSBEF 196
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHRTATGLRKVVRELFTNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHRTATGLRKVVRELFTNGARKNAFKILVI 256
QY 241 TDGEXFGDPLGYEDVIPLEDRGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDGEXFGDPLGYEDVIPLEDRGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHVFQIN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAAITSNGLPLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGQTIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEQTRGGQSVQVCL 480
DB 437 QNTGMWESNANVKGQTIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEQTRGGQSVQVCL 496
QY 481 PRGQARWQCDVLYGEOQPKGRFGAALTVLGDVNGDKLTDAI GAFGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEOQPKGRFGAALTVLGDVNGDKLTDAI GAFGEEDNRGAVYLF 556
QY 541 HGTSGSGI:SPHSQRIAGSKLSPRLQYFGQSLGGDLTMDGLVDTLVGAQGHVLLRSQ 600
DB 557 HGTSGSGI:SPHSQRIAGSKLSPRLQYFGQSLGGDLTMDGLVDTLVGAQGHVLLRSQ 616
QY 601 PVLVRKAIWFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLVRKAIWFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQPLNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQPLNCIEDPVSPIVLRNF 736

QY 721 SLVGTPLSAFQNLRPVLAEDAQRFTALFPPEKKNCGNDNI CODDLISITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFQNLRPVLAEDAQRFTALFPPEKKNCGNDNI CODDLISITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNGEDSYRTQVTFPPFDLSYRKVSTLQNRQSRWSRLACESASSTEV 840
DB 797 GPREFNVTVVRNGEDSYRTQVTFPPFDLSYRKVSTLQNRQSRWSRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIFFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
QY 901 QLELPKVAVYVMTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV 960
DB 917 QLELPKVAVYVMTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV 976
QY 961 RLNQTVIWDROPVTFSENLSTCHTKERLPSHSDFLAELKAPVNVVNCISIAVCQRIQCDDIP 1020
DB 977 RLNQTVIWDROPVTFSENLSTCHTKERLPSHSDFLAELKAPVNVVNCISIAVCQRIQCDDIP 1036
QY 1021 PFGIQEEFNATLKGNLSPDWYIKTSHNHLLIVSTAEILLFNDVSFTLLPGQCAFVRSQTET 1080
DB 1037 PFGIQEEFNATLKGNLSPDWYIKTSHNHLLIVSTAEILLFNDVSFTLLPGQCAFVRSQTET 1096
QY 1081 KVEPFEVNPPLIVGSSVGGLLLALLITALYKLGFFKROYKDMMSGGPPGAEPPQ 1137
DB 1097 KVEPFEVNPPLIVGSSVGGLLLALLITALYKLGFFKROYKDMMSGGPPGAEPPQ 1153
RESULT 5
AA014428
ID AA014428 standard; protein; 1153 AA.
AC AA014428;
XX
XX 03-MAY-2002 (first entry)
XX Integrin Mac-1 alpha subunit.
XX
XX Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
XX open conformation; integrin related inflammatory disorder;
XX integrin related immunological disorder; rheumatoid arthritis; ischaemia;
XX reperfusion; hypovolemic shock; infarction; cerebral shock;
XX viral infection; cancer; gene therapy; vaccine;
XX bioactive agent screening.
OS Unidentified.
XX
XX WO200204521-A2.
XX
XX 17-JAN-2002.
XX
XX 09-JUL-2001; 2001WO-US21805.
XX
XX 07-JUL-2000; 2000US-216600P.
XX
XX (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
XX (BLOO-) CENT BLOOD RES.
XX
XX Springer T;
XX
XX WPI; 2002-148167/19.
XX
XX New integrin I domain protein having alteration in at least 2
XX noncontiguous regions and exits in an open conformation, useful for
XX treating, preventing or suppressing inflammatory or immunological
XX disorders
XX
XX Example 1; Fig 1F; 90pp; English.
PS
XX The invention comprises structurally biased variant integrin inserted (I)
XX domain proteins, wherein the alterations to the protein occur in at least
CC

two noncontinuous regions. Specifically the the variant integrin I domain proteins are structurally biased to exist in the open conformation, thereby altering the binding ability of the protein. The invention also comprises nucleic acids encoding the variant integrin I domain proteins. The integrin I domain proteins and nucleic acids are useful for treating, preventing or suppressing integrin related inflammatory and immunological disorders (e.g. rheumatoid arthritis). The variant integrin I domain proteins and nucleic acids can also be used for treating: ischaemia/ reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral infection; and cancer. The variant integrin I domain nucleic acids and proteins may be used in gene therapy, as vaccines and to screen for bioactive agents. The present amino acid sequence represents the Mac-1 alpha subunit of integrin.

SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 23; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFQSVQVQLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTQENARGFQSVQVQLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLPFLGSLNRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLPFLGSLNRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKEMVSTVMEQLKKSTLPSLMQYSEEP 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKEMVSTVMEQLKKSTLPSLMQYSEEP 196

QY 191 RIHFTFKFQNNPRLSKPTTOLLGRTHATGLRVRELVNITNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPRLSKPTTOLLGRTHATGLRVRELVNITNGARKNAFKILVVI 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKQELNTVASKPRDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKQELNTVASKPRDRHVFQIN 316

QY 301 NFEALKTQNLREKIFAIEGTQGTSSSPFHEMSQEGFSAATISNGPLLTGVSQYDWAG 360
DB 317 NFEALKTQNLREKIFAIEGTQGTSSSPFHEMSQEGFSAATISNGPLLTGVSQYDWAG 376

QY 361 GVFLYTSKESTFFINMTVDSDMDAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKESTFFINMTVDSDMDAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLCAPHYEYQTRGGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLCAPHYEYQTRGGQVSVCP 496

QY 481 PRGQARWQCDVAVLGEQGPWGRFGAALTVLGDVNGDKLTDAIGAAGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLGEQGPWGRFGAALTVLGDVNGDKLTDAIGAAGEEDNRGAVYLF 556

QY 541 HCTSGSGISPSHSORIASKLSPLRYQGSISGGQDLTMQGLVDLTVGAQGHVLLRSQ 600
DB 557 HCTSGSGISPSHSORIASKLSPLRYQGSISGGQDLTMQGLVDLTVGAQGHVLLRSQ 616

QY 601 PVLVRKAIEMFNPVARNVFCNDQVWKGKAGEVRVCLHVQKSTRDLRREGQTSVVT 660
DB 617 PVLVRKAIEMFNPVARNVFCNDQVWKGKAGEVRVCLHVQKSTRDLRREGQTSVVT 676

QY 661 YDLALDGRPHSRVFNFTKSTRQTVGLTQTCTELKQLPNCIEDPVSPIVLRNLF 720
DB 677 YDLALDGRPHSRVFNFTKSTRQTVGLTQTCTELKQLPNCIEDPVSPIVLRNLF 736

QY 721 SLVGTPLSAFGLNRLPVLAEADQRLFTALFPPEKNCNDNI CODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGLNRLPVLAEADQRLFTALFPPEKNCNDNI CODDLSITFSFMSLCLVVG 796

QY 781 GPREFNVTVVRNDGSDSYRTQVTFPFLDLISYRKVSTLQNRQSRWRLACESASSTEV 840

DB 797 GPREFNVTVVRNDGSDSYRTQVTFPFLDLISYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANTVSENMPRTNKTPE 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANTVSENMPRTNKTPE 916
QY 901 QLELPVKYAVYVMTVTSHGVSSTKYLNFTASENTSRVMQHVQVSNLQSRSLPSLVLPVPV 960
DB 917 QLELPVKYAVYVMTVTSHGVSSTKYLNFTASENTSRVMQHVQVSNLQSRSLPSLVLPVPV 976

QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDIP 1020
DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDIP 1036

QY 1021 FFGIQSEFNATLKGNSLFDWYIKTSHNHLILVSTABILFNDSVFTLLPGOGAFVRSOTET 1080
DB 1037 FFGIQSEFNATLKGNSLFDWYIKTSHNHLILVSTABILFNDSVFTLLPGOGAFVRSOTET 1096

QY 1081 KVEPPEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPPQ 1137
DB 1097 KVEPPEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGGPPGAEPPQ 1153

RESULT 6
AAR04136
ID AAR04136 standard; protein; 1153 AA.
XX AC AAR04136;
XX DT 25-MAR-2003 (updated)
DT 07-SEP-1990 (first entry)
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
XX non-specific defence system; integrin gene superfamily.
XX OS synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 86..88 /*label= putative N-glycosylation site
FT Modified-site 240..242 /*label= putative N-glycosylation site
FT Modified-site 391..393 /*label= putative N-glycosylation site
FT Modified-site 469..471 /*label= putative N-glycosylation site
FT Modified-site 693..695 /*label= putative N-glycosylation site
FT Modified-site 697..699 /*label= putative N-glycosylation site
FT Modified-site 735..737 /*label= putative N-glycosylation site
FT Modified-site 802..804 /*label= putative N-glycosylation site
FT Modified-site 881..883 /*label= putative N-glycosylation site
FT Modified-site 901..903 /*label= putative N-glycosylation site
FT Modified-site 912..914 /*label= putative N-glycosylation site
FT Modified-site 941..943 /*label= putative N-glycosylation site
FT Modified-site 947..949 /*label= putative N-glycosylation site
FT Modified-site 979..981 /*label= putative N-glycosylation site
FT Modified-site 994..996 /*label= putative N-glycosylation site
FT Modified-site 1022..1024 /*label= putative N-glycosylation site
FT Modified-site /*label= putative N-glycosylation site

FT Modified-site 1045..1047
 FT /*label= putative N-glycosylation site
 FT Modified-site 1051..1053
 FT /*label= putative N-glycosylation site
 FT Modified-site 1076..1078
 FT /*label= putative N-glycosylation site
 FT region 1..16
 FT /*label= signal peptide
 FT region 1106..1134
 FT /*label= putative transmembrane region
 XX
 PN EP364690-A.
 XX
 PD 25-APR-1990.
 XX
 PF 17-AUG-1989; 89EP-0115159.
 XX
 PR 23-AUG-1988; 88US-0235353.
 PR 09-MAR-1989; 89US-0321239.
 XX
 XX (DAND) DANA FARBER CANCER INST INC.
 XX
 XX
 PI Springer TA, Corbi A;
 XX WPI; 1990-125938/17.
 DR N-PSDB; AAQ04043.
 DR
 XX
 PT New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
 PT inflammation and viral infections, and in diagnosis
 XX
 PS Disclosure; Page 7; ?pp; English.
 XX
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue.
 CC Mac-1 is a member of the Integrin Gene superfamily.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1153 AA;
 Query Match 99.3%; Score 5842; DB 11; Length 1153;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1127; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVWVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVWVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSIGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
 DB 77 RLQVPVEAVNMSIGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136
 QY 121 FPEALRGCEQSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQKKSKTLFSLMQYSEEF 180
 DB 137 FPEALRGCEQSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQKKSKTLFSLMQYSEEF 196
 QY 181 RIHFTKFEQNNPNRSLKIPITQLGRTHATGKRVKVVRELFNITNGARKNAKIFLFL 240
 DB 197 RIHFTKFEQNNPNRSLKIPITQLGRTHATGKRVKVVRELFNITNGARKNAKIFLVI 256
 QY 241 TQGEKFGDPLGYEDVPELDRGVIRYVIGVDGAPRSEKSRQELNTVASKPRDHVQFN 300
 DB 257 TQGEKFGDPLGYEDVPELDRGVIRYVIGVDGAPRSEKSRQELNTVASKPRDHVQFN 316
 QY 301 NFEALKTIQNLREKIFAIEGTQTSSEFHEMSQEGFSAATSNGLLSTVGSVDWAG 360
 DB 317 NFEALKTIQNLREKIFAIEGTQTSSEFHEMSQEGFSAATSNGLLSTVGSVDWAG 376
 QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
 DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLICAPHYYEQTRGGQVSVCP 480
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLICAPHYYEQTRGGQVSVCP 496
 QY 481 PRGQARWQCDVLYGEGQGWGRFGAALTVDVGVNGDKLTVDVGAIGAPGEDNRGAVYLF 540
 DB 497 PRGQARWQCDVLYGEGQGWGRFGAALTVDVGVNGDKLTVDVGAIGAPGEDNRGAVYLF 556
 QY 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVDVGAIGHVLLRSQ 600
 DB 557 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVDVGAIGHVLLRSQ 616
 QY 601 PVLRVKAIEMFNPREVARNFECNDQVYKGEAGEVRVCLHVQKSTRDLREGQTQSVVT 660
 DB 617 PVLRVKAIEMFNPREVARNFECNDQVYKGEAGEVRVCLHVQKSTRDLREGQTQSVVT 676
 QY 661 YDLALDSGRPHSRVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
 DB 677 YDLALDSGRPHSRVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736
 QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
 DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900
 DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 916
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPLISLVFLVPV 960
 DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPLISLVFLVPV 976
 QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVNVNCIAVCQRIQCDDIP 1020
 DB 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVNVNCIAVCQRIQCDDIP 1036
 QY 1021 PFGIOEENATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQCAFVRSQTET 1080
 DB 1037 PFGIOEENATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQCAFVRSQTET 1096
 QY 1081 KVEPFEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPQ 1137
 DB 1097 KVEPFEVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPQ 1153
 RESULT 7
 AAR07120
 ID AAR07120 standard; protein; 1163 AA.
 XX
 AC AAR07120;
 XX
 DT 25-MAR-2003 (updated)
 DT 05-FEB-1991 (first entry)
 XX
 XX p150.95 alpha subunit encoded by clone lambdaX47.
 DE
 XX p150, 95 leucocyte adhesion receptor alpha-subunit;
 KW hairy cell leukaemia, rhinovirus.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 61..63
 FT /label-glycosylation site
 FT Modified-site 89..91
 FT /label-glycosylation site
 FT Modified-site 385..387
 FT /label-glycosylation site

DE Human Beta-integrin CD11c subunit protein.
 XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
 KW rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX US5728533-A.
 XX 17-MAR-1998.
 XX 07-JUN-1995; 95US-0485618.
 XX 07-JUN-1995; 95US-0485618.
 PR 23-DEC-1993; 93US-0173497.
 PR 05-AUG-1994; 94US-0286889.
 PR 21-DEC-1994; 94US-0362852.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van DER VIEREN M;
 XX WPI; 1998-206565/18.
 XX Screening assay for modulators of integrin binding - using
 PT immobilised or labelled alpha-d polypeptide, useful for, e.g.
 PT treating type-I diabetes
 XX Example 5; Fig 1A-D; 106pp; English.
 XX This sequence represents a human beta-integrin CD11c subunit which is
 CC used to describe a method for identifying compounds that modulate the
 CC interaction of the beta-integrin alpha-d subunit with a binding partner
 CC of alpha-d which involves contacting an alpha-d polypeptide with an
 CC alpha-d binding partner, one of which is immobilised and the other of
 CC which is labelled, in the presence of a test compound, and determining if
 CC the compound affects binding between the alpha-d polypeptide and alpha-d
 CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
 CC comprising the cytoplasmic, transmembrane or extracellular domain of
 CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
 CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
 CC asthma, psoriasis, lung inflammation, acute respiratory distress
 CC syndrome and rheumatoid arthritis.
 XX Sequence 1163 AA;
 SQ Query Match 58.4%; Score 3436; DB 19; Length 1163;
 Best Local Similarity 60.6%; Pred. No. 4.1e-278;
 Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;
 QY 1 ENLDTENAMTFOENARFGQSVVVOLOGSRVVGAPQEIIVANORGLYQCDYSTGCEPI 60
 DB 20 FNLDTBELTAFRVDSVGVQYANVWVVGAPQKIIAANQIGLYQCGYSTGACEPI 79
 QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTVVKGLCFGLGNSLRQPOK 120
 DB 80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHCEGRNWLTLGLFLGPT--QLTOR 137
 QY 121 PFEALRGCPQOESDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKSKTLFSLMQLVSEEF 180
 DB 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFAFMNFRVAVISQFQPSQFSLMQPSNKF 197
 QY 181 RHFTFEKQNNPNRSLKPIPTOLLGRTHRTATGLKRVRELNITNGARKNAFKILFL 240
 DB 198 QTHFTFEERFRTSNPLSLASVHLQGLQFTTATQNVVHRLFHASYGARRDAIKILVI 257
 QY 241 TGEKFGDPLGYEDVPELDREGVIRYVIGVDAPFSEKSRQELNVTASKPRDHVQFN 300
 DB 258 TCGKEGSDLDYKDVIPMDAGIIRYALGVGLAFQNRNSWKLNDIASKPQEHIFKVE 317
 QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAQ 360

DB 318 DFDALKDIONQKEKIFAIEGTETISSSFELEMAQEGFSAVFTPDGVLGAVGFTMSG 377
 QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRNOSLVLCAPRYOHIGLVAMFR 420
 DB 378 GAFLYPPNMSPTFINNSQENVMDRDSYLGYSTELALWKGVQSLVLCAPRYOHIGKAVFI 437
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYYEQTRGGQSVCP 480
 DB 438 QVSRQWRMAEVIQIGSYFGASLCSDVDVDSNGSTDVLVIGAPHYYEQTRGGQSVCP 497
 QY 481 PRGQARWQCDAVLYGEOQPMGRGAALTIVLDVNGDKLTDTVAIGAGEENRGAAYLF 540
 DB 498 PRGWR-RWMCDAVLYGEOQPMGRGAALTIVLDVNGDKLTDTVAIGAGEENRGAAYLF 556
 QY 541 HGTSGGISPSHSQRIAGSKLSPRLQYFGQSLSGQDITMDGLVLTGVAQGHVLLRSQ 600
 DB 557 HGVLGPSISPSHSQRIAGSKLSSRLQYFGQSLSGQDITMDGLVLDVAVGARQVLLRTR 616
 QY 601 PVLRYKAIMFNPREVARNFECNDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
 DB 617 PVLWVGSMQFIPAEIPRSAFECEQVVSQTLVQSNICLYDKSKNLLSGSRDLQSSVT 676
 QY 661 YDLALDSCRPHSRVAFNETKNSRTOVGLGTQTCETLKLQLPNCIEDPVSIVLRNF 720
 DB 677 LDALAPGLSPRAIFQETKNSRVRVGLGKAHCENFNLLPSCVEDSVIPIILRNF 736
 QY 721 SLVGTPLSAFAGNLRPLAEDAQRLFTALFPPEKNCNDNICODLSITFSFMSLCLVVG 780
 DB 737 TLVGKPLLAFLNLRPLAALAQRYFTASLPFEKNCADHICODNLGISFSFPLGKSLVG 796
 QY 781 GPREFNVTVRNDEGDSYRTQVTFPFPLDLSYRKVSTLQNRQSRQWLACASASSTEV 840
 DB 797 SNLELNAEVMVWMDGEDSYGTTITFSHPAGLSYRVAEGQKQLRSLHLTC--CSAPVG 854
 QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
 DB 855 SQGTWSTSCRINHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENNIPRTSKTIF 914
 QY 901 QLELPKVAVVMVTSHGVSSTYKLNFTAS-ENTSRVMOHOYOVNSLQORSLSILVFLVP 959
 DB 915 QLELPKVAVIVVSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSINFWVP 974
 QY 960 VRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQCDI 1019
 DB 975 VELNQEAVMDEVESHQPNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIACLRFCDV 1034
 QY 1020 PFFGQOEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTE 1079
 DB 1035 PFSVQOELDFTLKGNLSFGWVRQTLQKKVSVSVAEIIFDTSVYSQLPQGEAFMRAQTI 1094
 QY 1080 TKVEPFEVNPPLVLTGSSVGGLLLLALITLALYKLGFFKQYKDMSE 1128
 DB 1095 TVLEKYKVNPIPLVIGSSIGGLLLALITVLYKVGFFKQYKEMMBE 1143
 RESULT 9
 AAB07361
 ID AAB07361 standard; Protein; 1163 AA.
 XX
 AC AAB07361;
 XX
 DT 17-JAN-2001 (first entry)
 XX
 DE Human CD11c protein sequence.
 XX
 KW Human; macrophage infiltration inhibition; alpha_d integrin;
 KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
 KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
 KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
 KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
 KW rheumatoid arthritis; central nervous system injury; CD11c.
 XX

OS Homo sapiens.
 XX WO200029446-A1.
 XX 25-MAY-2000.
 PD 16-NOV-1999; 99WO-US27139.
 XX 16-NOV-1998; 98US-0193043.
 PR 08-JUL-1999; 99US-0350259.
 XX (ICOS-) ICOS CORP.
 PA Gallatin MW, Van Der Vieren M;
 XX WPI; 2000-387751/33.
 XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous
 PT system injury sites
 XX Example 5; Fig 1; 270pp; English.
 PS Integrins are a class of membrane-associated molecules that participate
 XX in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit; the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11c. This
 CC sequence was used in an alignment to identify a novel beta2 integrin
 CC alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has
 CC approximately 66% identity to the protein sequence of alpha d. The
 CC Alpha d gene and protein may be useful in therapy for diseases linked
 CC to aberrant alpha d function (e.g. Type 1 diabetes, atherosclerosis,
 CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
 CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
 CC adhesion deficiency (LAD). In addition, anti-alpha d monoclonal
 CC antibodies may be used in the inhibition of macrophage infiltration at
 CC the site of a central nervous system injury. The monoclonal antibodies
 CC can also be used to detect and diagnose Crohn's disease.
 XX
 SQ Sequence 1163 AA;
 Query Match 58.4%; Score 3436; DB 21; Length 1163;
 Best Local Similarity 60.6%; Pred No. 4, 1e-278;
 Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;
 QY 1 FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQEIIVAAANQKSLVQCDYSTGSCPEI 60
 DB 20 FNLDTEELTAFRVDSAGDSVVQVANSWVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79
 QY 61 RLOVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVGKLCFLFGSNLRQOPK 120
 DB 80 GLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVGKLCFLFGSNLRQOPK 137
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTFLSLMOYSEEF 180
 DB 138 LPVSRQECPRQEQDVLFLIDGSGSISRNFATMFWRAVISQFQRPSTQSLMQPSNKF 197
 QY 181 RIHTFKFQNNPRSLKIPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILPLL 240
 DB 198 QHTFTFEFRRTSNPLSLASVHQLQGGTYTATAIQNVVHRLFHASYGARRDAIKILVI 257
 QY 241 TDGEKFGDPLGVEDVPELDREGVTRYVIGVDFAFRSKSQELNTVASKPRDRHVFOIN 300
 DB 258 TDGKKEGSLDYKDVIPNADAGIIRYAIQVGLAFQNRNSWKELNDIASKPSQEHFKVE 317
 QY 301 NFEALKTQNLQREKIFAETGTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
 DB 318 DFDAUKDQNLQKEKIFAETGTETISSSSFELEMAQEGFSAVFTPDGVLGAVGVSFTWSG 377
 QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVILGAPRYQHIGLVAMFR 420

DB 378 GAPLYPPNMSPTFFINNSQENVDNRUSYLGYSYTELAWKGVQSLVLCAPRYQHIGKAVIFI 437
 QY 421 QNTGMWESNANVKGTOIGAYFCASLCSVDSDNGSTDLVLI GAPHYEQTGGQSVSCPL 480
 DB 438 QVSRQWRMKAIEVIGTIGSYFCASLCSVDSDNGSTDLVLI GAPHYEQTGGQSVSCPL 497
 QY 481 PRGQARWOCDAVLYGEGQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEENRGAVYLF 540
 DB 498 PRGWR-RWMCDAVLYGEGQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEENRGAVYLF 556
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSISGGQDLTMDGLVDLTVCAOQHVLRLRSG 600
 DB 557 HGVLGPSISPSHSQRIAGSKLSPRLQYFGQSISGGQDLTMDGLVDLTVCAOQHVLRLRSG 616
 QY 601 PVLVRKAIMFNPREVARNVFECDNQVKKEAGEVVRVCLHVQKSTRDLRERGOIQSVVT 660
 DB 617 PVLWGVSMQFIPAEIPRSFAFECREQVSEQTLVQSNICLYIDKRSKNLGRDLQSSVT 676
 QY 661 YDLALDSGRPHRAVFNETKSTRQTVLGLTQTCETLKLQPLNCIEDPVSPIVRLNPF 720
 DB 677 LDALAPGRSLSPRAIFQETKRSLSRVRLGLKAHCENFNLLPSCVSDVSPVPIILRLNF 736
 QY 721 SLVGTPLSAFGLNLRPVLAEDAORLFTALPFEKNKCGNDNICDDLSITTFSEMSLDCLVVG 780
 DB 737 TLVGKPELLAFNLRPMLAALAQRYFTASLPFEKNKCGADHICQNLGISFSFPGKSLVVG 796
 QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVYSTLQNRQSRQSRMRLACESASSTEV 840
 DB 797 SNLELNAEVMVWMDGEDSYRTQVTFPPFLDLSYRKVYSTLQNRQSRQSRMRLACESASSTEV 854
 QY 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
 DB 855 SQGTWSTSCRINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 914
 QY 901 QLELPVKYAVVMVVTSHGVSTKVLNFTAS-ENTSRVMOHYOVSNLQSRSLPSISLVFLVP 959
 DB 915 QLELPVKYAVVMVVTSHGVSTKVLNFTAS-ENTSRVMOHYOVSNLQSRSLPSISLVFLVP 974
 QY 960 VRLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI 1019
 DB 975 VELNQEAVMDVEVSHVHPQNPQLSCSEKTIAPPASDFLAHQKPNVLDCSIAGCLRPRCDV 1034
 QY 1020 PFGIOQEFNATLKGMLSPDWYIKTSHNHLLIIVSTAEILFNDVSFTLLPQOGAFVRSQTE 1079
 DB 1035 PSFSVQBELDFTLKGMLSPDWYIKTSHNHLLIIVSTAEILFNDVSFTLLPQOGAFVRSQTE 1094
 QY 1080 TKVEPEVPNPPLPIVGVSSVGGLLALITAAALYKLGFFKQYKDMWSE 1128
 DB 1095 TVLEKYKVNPIPLIVGVSSVGGLLALITAAALYKLGFFKQYKDMWSE 1143
 RESULT 10
 ABG61470
 ID ABG61470 standard; Protein; 1163 AA.
 XX AC ABG61470;
 XX AC AC
 DT 27-AUG-2002 (first entry)
 XX Human Beta2 integrin alphaCD11c subunit.
 DE Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
 XX LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 XX autonomic dysfunction; sensory dysfunction; spinal cord injury.
 OS Homo sapiens.
 XX

PN W0200230980-A2.
 XX 18-APR-2002.
 XX 15-OCT-2001; 2001WO-US32059.
 XX 13-OCT-2000; 2000US-0688307.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van Der Vieren M;
 PI WPI; 2002-463260/49.
 XX
 XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
 PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
 PT or limiting autonomic and sensory dysfunction following spinal cord
 PT injury.
 XX
 XX Example 5; Page 194-198; 270pp; English.
 XX
 CC The invention relates to promoting locomotor recovery, inhibiting
 CC locomotor damage, limiting locomotor impairment, or limiting autonomic
 CC and sensory dysfunction following spinal cord injury by administering an
 CC anti-alpha-d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
 CC spinal cord injury victim. The method also involves the use of
 CC a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion
 CC molecule, vascular cell adhesion molecule). The method is useful for
 CC promoting locomotor recovery, inhibiting locomotor damage, limiting
 CC locomotor impairment, or limiting autonomic and sensory dysfunction
 CC following spinal cord injury. In particular, the spinal cord injury
 CC comprises compression of the spinal cord. The antibodies are also useful
 CC for reducing inflammation at the site of a central nervous system injury.
 CC The specification also details the identification of Beta2 integrin
 CC alphaD cDNAs and proteins, for use in raising the antibodies. Beta2
 CC integrins are implicated in diseases such as LAD (leukocyte adhesion
 CC deficiency, inflammatory response, diabetes, multiple sclerosis,
 CC arthritis, graft atherosclerosis, inflammatory bowel disease,
 CC Crohn's disease, ulcerative colitis, immune complex alveolitis
 CC and leukaemia. The present sequence is a Beta2 integrin alpha subunit
 CC sequence included for comparison with the Beta2 integrin alphaD protein
 CC sequences.
 XX
 XX Sequence 1163 AA;
 SQ

Query Match 58.4%; Score 3436; DB 23; Length 1163;
 Best Local Similarity 60.6%; Pred. No. 4.1e-278;
 Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGFGQSVVQLGSGRRVVGAPQEIIVAAANQSGLYQCDYSTGSCPEI 60
 DB 20 FNLDTEELTAFRVDSAGFGDSVVQYANWVVGAPQIIAANQIGGLYQCGYSTGACEPI 79
 QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQCSTENTYKVGCLFELGSLNLRQPOK 120
 DB 80 GLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHHECGRNMYLTGLCLLGP--QLTOR 137
 QY 121 FPEALRGCPQEDSDIAPLDIGSGSIIPHDPRRMEKWSVMEQLKSKSLFSLMOYSEEF 180
 DB 138 LPVSRQECPRQEQDIIVFLDIGSGSISRNFATMNFVRAVISQFQPSQFSLMQSNNKF 197
 QY 181 RIHFTKPEQNNPNRSLKIPITQLGRTHATGLRKVVRELFNITNGARKNAKFLIFLL 240
 DB 198 QTHFTPEEPRTSNPLSLASVHQLQGFYVTAIQNVRHLPHASVGARRDAIKILIVI 257
 QY 241 TDGEKGDPLGEDVITPELDREGVIRYVIGVDAPRSEKSRQBLNTVASKPRPDHFQIN 300
 DB 258 TDGKEGDSLDYKDVIPMDAAGIIRIYVIGVLAQFNRRNSWKEINDIASKPSQEHIFKVE 317
 QY 301 NFEALKTIONQREKIFAIEGTOTGSSGFHEHMSQEGFSAAITTSNGPLLSVTVGSDWAG 360
 DB 318 DFDALNDIQNLKEKIFAIEGTETISSSFELEMAQEGFSAVTPDGPVLGAVGSTWSG 377

QY 361 GVFLYTSKEKSTFIMNTRVDSMDNDAYLQYAAAIILNRNVSQSLVLCAPYQHIGLVAMFR 420
 DB 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALWKGVSQSLVLCAPYQHIGKAVIFI 437
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYETOTRGQVSVCP 480
 DB 438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDSNGSDTLVLIGAPHYETOTRGQVSVCP 497
 QY 481 PRQQRARMQCDVLYXGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEENRGAVYLF 540
 DB 498 PRGWR-RMWCDVLYXGEOQPMGRFGAALTVLGDVNGDKLTDVWIGAPGEENRGAVYLF 556
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLOYFQOSLSGGQDLTMDGLVDLTVAQGHVLLLRQ 600
 DB 557 HGVLPSPISPSHSQRIAGSKLSPRLOYFQOSLSGGQDLTMDGLVDLTVAQGHVLLLRQ 616
 QY 601 PVLRLVKAIMFENPREVARNVFECNDQVVKGEVVRCLHVQKSTRDLREGQIQSVVT 660
 DB 617 PVLWVGVSQMFIPAPISAFECRQOVSEQTLVQSNICLYIDKRSKNLGSRDQSSVT 676
 QY 661 YDLALDSGRPHSRVFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
 DB 677 LDALAPGRLSPRAIFQETKNRSLSRVRVLGLKAHCENFNLLPSCVEDSVIPIILRNF 736
 QY 721 SLVGTPLSAFGNLRPVLADAOPLTALPPFPKNCNDNICQDDLSITFSFMSLDCLVVG 780
 DB 737 TLVGKPLLAFLNRLPMLAALAOQRYTASLPFFKNCADHICQDNLGISFSPGLKSLVG 796
 QY 781 GPREFNVTVTVNDGEDSVRTQVTFPFLDLSYRVKSTLQNSORSORSWBLACESASSTEV 840
 DB 797 SNLELNAEVMWVNDGEDSVGTTITTSHPAGLSYRVABGQKQOQKLSHLTC--CSAPVG 854
 QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTTF 900
 DB 855 SQGTWSTSCRIHNLIFRGGAQITFLATFVSPKAVCLDRLLLIJANVSENNIPRTSKTIF 914
 QY 901 QLELPKVAYVWVTVSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLQGRSLPISLVFLVP 959
 DB 915 QLELPKVAYVTVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQGRDLFVSNFWVP 974
 QY 960 VRLNQTIVWDRPQVTFSENLSSTCHKERLPSHSDFLAELRKAPVNVCSIANVQRCODI 1019
 DB 975 VELNQEAVMVDVEVSHPPQNPSCSEKIAPPASDPLAHIQKNPVLDCSIAGCLRCDV 1034
 QY 1020 PFFGIQEEFNATLKGNSLDFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQGAFFRSQTE 1079
 DB 1035 PSFSVQEEELDFTLKGNLSFGWVRQILQKKVSVSVVAEIIFDTSVYSQLPQENAFRAQTI 1094
 QY 1080 TKVEPFEVDPNPLIVGSSVGGLLLLALITALYKLGFPFKROYKQMMSE 1128
 DB 1095 TVLEKYKVHNPILIVGSSIGGLLLALITAVLYKVGFPPKROYKEMMEE 1143

RESULT 11
 AB007406
 ID AB007406 standard; Protein; 1163 AA.
 XX
 AC AB007406;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Protein differentially regulated in prostate cancer #9.
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
 XX
 OS Homo sapiens.
 XX
 PN W0200281638-A2.
 XX
 PD 17-OCT-2002.
 XX

PF 08-APR-2002; 2002WO-US10824.
XX
PR 06-APR-2001; 2001US-281731P.
PR 06-APR-2001; 2001US-281732P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Jay G;
XX
XX WPI; 2003-058520/05.
DR
XX Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
PT patients -
XX
PS Claim 1; Page 225-228; 416pp; English.
XX
CC The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.
XX
SQ Sequence 1163 AA;
Query Match 58.4%; Score 3434; DB 24; Length 1163;
Best Local Similarity 60.5%; Pred. No. 6.1e-278;
Matches 683; Conservative 140; Mismatches 300; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFQENARGFQSVVLQGSRRVVVVGAPQEIIVAAHQRSGLQCYDSTGSCPEI 60
DB 20 FNLDTEELTAFRVDSAGDSVVQVANSWVVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMLGLSLAATTPSPQLLACGPTVHTQCSNTYVVKGLCFGLSGNLRQOQPK 120
DB 80 GLQVPPEAVNMLGLSLASTTSPQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
QY 121 FPEALRGCPQESDIAFLIDGSGSIIPHDFFRMKEMWSTVMEQLKSKTLFSLMOYSEEF 180
DB 138 LPVSQBECPQESQDIFVILIDGSGSISSRNFAFMNFRVAVISQFQRPSTQFSLMOFSNKF 197
QY 181 RIHFTFKFQNNPNRSLKPTITQLIGRTHATGLRKVVRELFNITNGARKNAFKILFUL 240

RESULT 12

AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX
AC AAR78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX

198 QTHLTPEFRRTSNPLSLASVHQLQGGFTYTATAIQNVVHRLPHASVGARRDATKILIVI 257
241 TDGEKFGDPLGYEDVIPELDREGVIRVYIGVGDAFRSEKSRQELNLTVASKPPRDHVFQIN 300
258 TDGKKEGDTLDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKELNDIASQSEHIFKVE 317
301 NFEALXTIQNLREKIFAIEGTQTGSSSPHEHMSQEGFSAAITSNGLPLSTVGSVDWAG 360
318 DFDALXDIQTLREKIFPIEGTETSSSPSELEMAQEGFSAVFTDGPVLGAVGSEPTWSG 377
361 GVFLYTSKESKSTINMTRVDSNMNDAYLGAAAILLRNRVOSLVLCAPRVOHIGLVAMER 420
378 GAFLYPPNMSPTFINMSQENVDRDSYLGYSTELALWKGVOSLVLCAPRYOHTGKAVIFT 437
421 QNTGMWESNANVKGTOIGAFYGASLCSVDVDSNGSTDVLIGAPHYETQTRGGQVSCPL 480
438 QVSRQWRMKAETVGTQIGSYFGPGLCSVDVDSGSTDVLIGPPHYETQTRGAQVSCPL 497
481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNRGAYILF 540
498 PRGWR-RWMCDAVLYGEGHPWGRFGAALTVLGVNGDKLTDVVIGAPGEENRGAYILF 556
541 HGTSGSISPSHSQRTAGSKLSRLOYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
557 HGVLPSPISPSHSQRTAGSKLSRLOYFGQALSGQDLTQDGLVDLAVGARGQVLLLRTR 616
601 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLRVQKSTRDRREGIOQSVWT 660
617 PVLWVGVSQMFPAEIPRSFAFECREQVVSQTLVQSNICLYIDKSKNLLGSRDLQSSVT 676
661 YDLALDSGRPHSRVAFNETKXSTRQTVLGLTQCTETLKLQLPNCIEDVPVILRLNF 720
677 LDIALDPGLSPRATFOETKNRSLSRVVLGLKAHCENFNLLLPSCVEDSVPTITRLNF 736
721 SLVCTPLSARGNLRPVLAEDAORLTALPFEKNCNGDNICODDLSITFEFMSLDCLVVG 780
737 TLVGKPLLAFLRNLRLPMLAADAQRYFTASLPFEKNCNGADHICQDMLGIFSFPGLKSLVG 796
781 GPREFNVTVTRNDEGDSYRTQVTFPFPLDLVSRKVSTLQVRSQRSWRLACESASSTEV 840
797 SNLELNAEVMWMDGSDSYGTTTFSSHAGLSYRYVAEGQKQGLRSLHLTCDSPVG-- 854
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTPEF 900
855 SQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVLGDRLLLTANVSSENTPRTSKTTF 914
901 QLELPVKYAVVMVYTHGVSTKYLNFAS-ENTSRVMQHOYOVNSLQORSPLISLVFLVP 959
915 QLELPVKYAVVYTVSSSHQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSINFWP 974
960 VRLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPVWNCIAVCQRIQCDI 1019
975 VELNQEAVWMDVEVSLPQNPSLRCSSEKIAGPSDFLAHQKNPVLDCSIAGCLFRCDV 1034
1020 PFFGIQEFENATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGQAFVRSOTE 1079
1035 PSFSVQBELDPLTKGNLSFGWVRQILOKKVSWSVAEITFTDTSVYSLPQGEAFMRAQTT 1094
1080 TKVEPFEVPNPLPLIIVGSSVGLLLALLITAAALYKLGFFKQYKDMSE 1128
1095 TVLEKYKVHNPTPLIVGSSIGGLLLALLITAVLYKVGFFRQYKEMMEE 1143

Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis; inflammatory bowel disease; asthma.
Homo sapiens.
Key Location/Qualifiers
Domain 17..1108
Region 150..352
/note= "this region is homologous to the insert common to CH1a,b,c and may be a site for interaction with ICAM family proteins"
Binding-site 465..474
/note= "putative cation binding site"
Binding-site 518..527
/note= "putative cation binding site"
Binding-site 592..600
/note= "putative cation binding site"
Region 1109..1128
Domain 1129..1161
/note= "transmembrane region"
Domain 1129..1161
/note= "cytoplasmic domain"
W09517412-A1.
PN
XX
XX
PD 29-JUN-1995.
XX
PF 21-DEC-1994; 94WO-US14832.
XX
PR 05-AUG-1994; 94US-0286889.
PR 23-DEC-1993; 93US-0173497.
XX
PA (ICOS-) ICOS CORP.
XX
XX
PI Gallatin WM, Van Der Vieren M;
DR WPI; 1995-240603/31.
DR N-PSDB; AAQ91712.
XX
PT Alpha sub-unit polypeptide of human beta 2 integrin - used to identify potential antiinflammatory agents, for the treatment of graft arteriosclerosis, inflammatory bowel disease, asthma, etc.
PS Claim 7; Page 82-87; 172pp; English.
XX
XX A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha-Tm1 was used to screen a human spleen cDNA library to identify clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d.
CC The cDNA was manipulated to allow expression of recombinant alpha-d subunit in COS and CHO cells.
XX
SQ Sequence 1161 AA;
Query Match 57.8%; Score 3401; DB 16; Length 1161;
Best Local Similarity 59.3%; Pred. No. 3.5e-275;
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
Qy 1 FNLDENAMTFOENARGGQSVWOLQGRVVVGAPOEIVAAORGLYQCDYSTGSCPEI 60
Db 17 FNLDVEPTFOEDAGFGQSVWQGGSRVVGAPELVAAOGRVLDCAATGMCPPI 76
Qy 61 RLQVPEAVNMSLGLSLAATSPQLLACGPTVHQTCTSENTYVKGCLCFLFGSNLRQOPQK 120
Db 77 PLHIRPEAVNMSLGLTAAASNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-EIIQT 135
Qy 121 FPEALRGCPQSDSIAFLIDSGSIIIPHDPRMKEWSTVNEQKSKTLFSLMQYSEEP 180
Db 136 VPDATPECPHOEMDIVELIDSGSIDQDNFQMKGFQVAVMGQFEGDTTLFALMQYSNLL 195
Qy 181 RIHTEFEFONNPRSLIKETITOLLGRTHATGLRKVRRELFNITNGARKNAFKILFL 240
Db 196 KIHFTFTQRTSPSQSLVDPIVQLKGLTFTATGILTVVTLQFHHKNGARKSANKILIVI 255

Qy 241 TDGEKFDGPIGYEDVPELDEGVIRVIVGVGAQPRSEKSRQELNVTASKPPRDRHVFOIN 300
Db 256 TDGQKYPLEYSVDVIPQAEKAGIIRYAIGVHAFOGPTARQELNVTASSAPPQDHVKVD 315
Qy 301 NFEALKTIONQREKIFAIEGTOTGSSSSPEHEMSQGFSAAITNSGPLSLTVGVSQWAG 360
Db 316 NFAALGSIQOLQEKIYAVEGTOSRASSPFQHEMSQGFSTALTMDGLFICAGVSPWSG 375
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAIAAILRNRVQSLVIGAPRYQHIGLVAMFR 420
Db 376 GAFLYPNMSPFTFNMSQENVMDRDSYLGYSYTELALWGVQNLVIGAPRYQHTGKAVFT 435
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVILGAPHYEYQTRGGQVSCPL 480
Db 436 QVSRQWRKAEVGTGTIGSYFGASLCSVDVDSNGSTDLVILGAPHYEYQTRGGQVSCPL 495
Qy 481 PRGORARWQCDVLYGEGOPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRYGLYF 540
Db 496 PRGORVQWQCDVLYGEGOPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRYGLYF 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 600
Db 556 HGASEGSIPSHSQRIASSQLSPRLOYFGQALSGGQDLTMDGLMDLAVGARGQVLLRSL 615
Qy 601 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVVRVCLHVQKSTRDRREGIOQSVT 660
Db 616 PVLKVGVMRFPVEVAKAVYRWEKPSALEAGDATVCLTIQKSLDQL--GDIQSSVR 673
Qy 661 YDLALDSGRPHSRVAVENETNSTRTQVGLTQTCETLKLQPNCLIEDVSPVILNLF 720
Db 674 FDLALDPRLTSSRAIFNETKNPTLTRKTLGLGHCHCETLKLUPDCVEDVVSPIILNLF 733
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALPFPEKNCNDNIQDDLSITFSFMSDCLVVG 780
Db 734 SLVREPIPSQNLRLPVLAGSQDLFTASLPFEKNCQDGLCEGDLGVTLSFGSLQTLTVG 793
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPDLVSRVYKSTLQNRQSRQSRWLACESASSTEV 840
Db 794 SSLELNVITVMNAGSDSYGVVSLYYPAGLSHRRVSGAQKPHQSALRLACETV-PTED 852
Qy 841 SGALKSTCSINHPPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 853 EG-LRRSRCSNVNHPHFEGSNGTFIVTFDVSYSKATLDGRLMLRASASSENKASSSKATF 911
Qy 901 QLELPVYAVYVMVVTSHGVSTKYLNFTASENTSRVMOHQVOVSNLQORSIPISLVFLVP 959
Db 912 QLELPVYAVYTMISRQESTKYNFATSDKWKKEAHEHYRVNVLNSQRLAISINFWP 971
Qy 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI 1019
Db 972 VLLNGVAVMDVMVMEAPSQL--PCVSRKPPQSHSDFLTQISRSPLMDCSIADCLQFRCDV 1029
Qy 1020 PFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGCGAFVRSQTE 1079
Db 1030 PSFSVQBELDFTLKNLSFGWVRETLQKKVYLVSVVAEITFDTSVYSQLPQGAQPMRAQME 1089
Qy 1080 TKVEPPEVPNPLPLVGVSSVGLLLALITAAALYKLGFFKRYQKDMWSE 1128
Db 1090 MWLEDEEVNAIFIIMGSSVGALLLALITATLYKLGFFKRYKEMLED 1138
RESULT 13
AAW23049
ID AAW23049 standard; Protein; 1161 AA.
XX
AC AAW23049;
XX
DT 24-FEB-1998 (first entry)
XX
DE Human beta 2 integrin alpha d subunit.
XX
KW Beta 2 integrin alpha d subunit; human; cell migration;
KW cell adhesion; phagocytosis; diabetes; atherosclerosis;

KW multiple sclerosis; aethma; psoriasis; lung inflammation;
 XX acute respiratory distress syndrome; rheumatoid arthritis.

OS Homo sapiens.

XX Key Location/Qualifiers

XX Domain 17..1108
 /label= Extracellular_domain

XX Domain 1109..1128
 /label= Transmembrane_domain

XX Domain 1129..1161
 /note= "homologous to the human CD11c transmembrane region"

XX Domain 150..352
 /label= Cytoplasmic_domain

XX Domain 150..352
 /note= "region homologous to the I (insertion) domain common to CD11a, CD11b and CD11c"

PN W09731099-A1.

XX 28-AUG-1997.

XX 24-FEB-1997; 97WO-US02713.

XX 22-FEB-1996; 96US-0605672.

XX (ICOS-) ICOS CORP.

XX Gallatin WM, Van Der Vieren M;

XX WPI; 1997-435154/40.

XX N-PSDB; AAT79220.

XX Hybridoma 199M and antibody secreted by it - specific for new rat
 beta2 integrin subunit, useful to detect subunit in cells and
 modulate its activity

XX Example 5; Page 116-120; 222pp; English.

CC This polypeptide comprises a novel human beta 2 integrin subunit,
 designated alpha d. Its sequence was deduced from a cDNA clone
 (see AAT79220) isolated from a spleen cDNA library. Alpha d is
 involved in cell migration, phagocytosis and cell-cell interaction.
 Recombinant alpha d polypeptides can be expressed in transformed
 host cells for use in assays for identifying antibodies or other
 compounds that modulate alpha d activity or which modulate the
 interaction between alpha d and a ligand, for treating or
 preventing diseases in which macrophages are implicated. Treatment
 is applicable to disease states in which alpha d binding, or
 localised accumulation of cells which express alpha d, is
 implicated such as type I diabetes, atherosclerosis,
 multiple sclerosis, aethma, psoriasis, lung inflammation, acute
 respiratory distress syndrome and rheumatoid arthritis.

SQ Sequence 1161 AA;

Query Match 57.8%; Score 3401; DB 18; Length 1161;

Best Local Similarity 59.3%; Pred. No. 3.5e-275;

Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFQENARGGVSVOLOGSRVVGCAPOEIVANORGSLQCYDSTGSCPEI 60

DB 17 FNLDVEEPTIFQEDAGGVSVOFGSRLVVGAPLEVAANTGRLYDCAATGMCQPI 76

QY 61 RLQVPVAVNNSLGLSAAATSPQLACGFTVHQTCSNTYVKGCLFFLGSNLRQOPQK 120

DB 77 PLHIRPEAVNNSLGLTAASTNGSRLAACGPTLHRVCGENSYSKSGCLLIGSRW-EIIQT 135

QY 121 FPALRCPPQDSDIAPLDGSSGIIPDPRKKEWSTVMEQLKSKTLFSLMQVSEEF 180

DB 136 VPATPCPCHEQMDIVFLIDGSGSIDQDNQMGKFVQAVMGQFEGDTDLFALMQYSNLL 195

QY 181 RIHFTKFEQNNPNRSLIKPITOLLGRTHATCLRKRVRELNIITNGARKNAFKILFL 240

DB 196 KIHFTTQFRTSPSQSLVDPIVQLKGLTFTATGTLTVVTLFHHKNGARKSAKILIVI 255
 QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDHFVQIN 300
 DB 256 TDGQKYKDPLEYSVDIPAQEKAGIIRYAIGVGHAFQGTARQELNTISSAPPQDPVFKVD 315
 QY 301 NFEALKTIQNLREKI PAIEGTOTGSSSSFEHEMSQEGSAAITSNGLPSTVGSYDNAG 360
 DB 316 NFAALGSIQKQLEKIYAVEGTQSRASSSPQHEMSQEGFSTALTMGDLFAGVGSFWSG 375
 QY 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLAVMR 420
 DB 376 GAFLYPPNMSPTFFINMSQENVDMRDSYLGYSYELALWKGQVNLVLGAPRYQHTGRAVFT 435
 QY 421 QNTGHWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLVIGAPHYQTRGGQVSVCPDL 480
 DB 436 QVSRQWRKKAQVETGTQIGSYFGASLCSDVDNSGSTDVLVIGAPHYQTRGGQVSVCPDL 495
 QY 481 PRGQARWQCDALVYGEQOPMGREGAALTVLVDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
 DB 496 PRGQVQWQCDALVYGEQOPMGREGAALTVLVDVNGDKLTDVAIGAPGEQENRGAVLYF 555
 QY 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDGLTMGDLVLTGAAQGHVLLRSQ 600
 DB 556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGQDGLTQDGLMDLAVGARGQVLLRSU 615
 QY 601 PVLRYKAIWERNPREVARNVFECNDQVWKGKAGEVRVCLVQKSTRDLRLEGGQTSVVT 660
 DB 616 PVLKVGAVAMRFPVEKAVRVCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRVFNETKSTRQTOVLGTQTCETIKLQLPNCIEDPVPVILRLNF 720
 DB 674 FDLALDPLRLTSRAIFNETKNTLTKRLKLGIIHCETIKLQLPNCIEDPVPVILRLNF 733
 QY 721 SLVGTPLSAFAGNLRLVLAEDAQRFLTALFPPEKNCNDNIIQDDLSITFSFMSLCLVVG 780
 DB 734 SLVREPIPSQNLRLVLAEGSODLFTASLPPEKNCQDGLCEGLGVTLFSGLQTLTVG 793
 QY 781 GRPENVTVTVNDGEDSYRTQVTFPPDLDSYKRVSTLQNRQSRQSRWLACESASSTEV 840
 DB 794 SSLELNVIITVMNAGEDSYGTVSVLYPAGLSHRRVSGAKQPHOSALACETV-PTED 852
 QY 841 SGALKSTCSINHPIFPENSEVTFNITPDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
 DB 853 EG-LKSSRCVNHPIFHEGSGNCTFIVTDPVSYKATLGDRLMLRASSENKASSSKATF 911
 QY 901 QLELPVKYAVVMVTSHGVSSTKYLNF-TASENTSRVMQHYQVSNLQGRSLPISLVFLVP 959
 DB 912 QLELPVKYAVYTMISRQEESTKYFNFATSDKKMKEAHRVYVNNLSQRDLAISINFWVP 971
 QY 960 VRLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQDI 1019
 DB 972 VLLNGVAVWVYMEAPSQSL--PCVSEKRPQHSDFLTQISRSPMLDCLSIADCLQPRCDV 1029
 QY 1020 PFEGIQEEFNATLKGNLSPDWIKTSHNLLIVSTAEILFNDSVFTLLPQCGAFVRSQTE 1079
 DB 1030 PSFSVQEEUDFTLKGNLSPGWVRETLQKVLVSVVAEITFTSVTSVQLPQGEAFRQAQME 1089
 QY 1080 TKVEPFEVNPPLPLIVSSVGGLLALITALYKLGFPFKRQYKMMSE 1128
 DB 1090 MVLEEDVYNAIPIIMSSVGALLLALITATLYKLGFPFKRHYKEMLED 1138

RESULT 14

AAW72825

ID AAW72825 standard; Protein; 1161 AA.

XX AAW72825;

AC AAW72825;

XX 19-JAN-1999 (first entry)

DE Human alpha-d.

XX Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
 KW immunoglobulin; fusion protein; binding molecule; antibody;
 KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
 KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
 KW acute respiratory distress syndrome;
 XX Homo sapiens.
 OS US5831029-A.
 PN 03-NOV-1998.
 XX 07-JUN-1995; 95US-0482293.
 XX 07-JUN-1995; 95US-0482293.
 PR 23-DEC-1993; 93US-0173497.
 PR 05-AUG-1994; 94US-0286889.
 PR 21-DEC-1994; 94US-0362652.
 XX (ICOS-) ICOS CORP.
 PA Gallatin WM, Van Der Vieren M;
 PI WPI; 1998-609318/51.
 XX N-PSDB; AAV67281.
 DR Antibodies specific for beta-2 integrin alpha-subunit d - useful in
 PT immunohistochemical analysis
 XX Example 5; Column 61-66; 106pp; English.
 XX The present sequence represents human alpha d. The present invention
 CC describes: (1) an antibody that specifically binds alpha d, which
 CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
 CC antibody as in (1); (3) an anti-idiotypic antibody specific for the
 CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
 CC monoclonal antibody of (2). Antibodies specific for alpha d can be
 CC used in immunohistochemical analysis to localise alpha d to subcellular
 CC compartments or individual cells within tissues. Substances that modulate
 CC alpha d binding (which may include antibodies) can be used to treat
 CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
 CC inflammation, acute respiratory distress syndrome or arthritis.
 XX Sequence 1161 AA;
 SQ Query Match 57.8%; Score 3401; DB 19; Length 1161;
 Best Local Similarity 59.3%; Pred. No. 3.5e-275;
 Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
 QY 1 FNLDENAMTQENARGQSVVQLQSGRVVVGAPQEIIVAAANORGLYOCDSYSGCEPI 60
 DB 17 FNLDVEEPTIQQDAGGQSVVQVQGGRLVVGAPLEVVAAANQTRLYDCAATGMCOPI 76
 QY 61 RLQPVVEANMSLGLSLAATSPOLLACGPTVHOTCSNTYVKGCLFGLSNLRQOPK 120
 DB 77 PLHIREAVNMSLGLTAASTNGSLACGPTLHRVCGENSYSKSCLLLSRW-EIIQT 135
 QY 121 FPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEWSTVMEQLKKSKTLFSLMQYSSEF 180
 DB 136 VPDATPECPHQEMDIVFLIDSGSIDQDNFMKGFVQAVMGQPEGTDTLFAIMQSNLL 195
 QY 181 RIHTEFKFQNNPNRSLIKPIQLLGRTHATGLRKVVRELFINITNGARKNAKILFL 240
 DB 196 KIHFTFTQFRSPSQSLVDPIVQLKGLTFTTGLTITVTLQFLHHKNGARKSAKKILVI 255
 QY 241 TDGEKFGPLGVEDVPELDREGVIRYVIGVDAPRSEKSRQELNTVASKPRDHVFOIN 300
 DB 256 TDGQYKDPLEYSDVIPQAEKAGIIRYAGVHAFQGTARQELNTISSAPPQDHVFKVD 315
 QY 301 NFELKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
 DB 316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGSFWSG 375

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMER 420
 DB 376 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSSTELALWKGQNLVLGAPRYQHTKAVIFT 435
 QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQSVCP 480
 DB 436 QVSQWRKKAETGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQSVCP 495
 QY 481 PRGORARWQCDVLYGEGOPWGRFGAALTIVLDGVNGDKLTDVAIGAPGEDNRGAYLYF 540
 DB 496 PRGORVQWQCDVLYRGEQHPWGRFGAALTIVLDGVNEDKLDVAIGAPGEDNRGAYLYF 555
 QY 541 HGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGGQDLTMDGLVDLTGAGQHVLRLSRQ 600
 DB 556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTODGLMDLAVGARGQVLLRLSL 615
 QY 601 PVLKVKAIMFNPREVARNVPECNDQVVKGEKEAGEVRLHVQKSTDRRLRREGQIQSVVT 660
 DB 616 PVLKVGVMARFSPVEVAKAVYRWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRAVNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLINLF 720
 DB 674 FDLALDPGLRTSRAIFNETKNPTLTRRKTIGLGIHCETLKLQLPDCVEDVVSPIILHLNF 733
 QY 721 SLVGTPLSAFQNLRLPVLAEADAQRLFTALFPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
 DB 734 SLVREPIPSQNLRLPVLAVGQDLFTASLPFEKNCNDGDLCEGDLGVTLSFGSLQTLTVG 793
 QY 781 GPREFNVTVVRNDGDSYRTQVTFPPDLSTVRKVTSTLQNRQSRQSRWLACESASTEV 840
 DB 794 SSLELNIVTVVRNAGDSYGTWVSLYYPAGLSHRRVSGAQKPHOSALRLACETV-PTED 852
 QY 841 SGALKSTSCSNHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
 DB 853 EG-LRSRCSNVHPIFHEGNGTFTVFDVSYKATLDGRMLRASASNNKASSKATF 911
 QY 901 QLELPVYAVYVMTSHGVSTKYLNF-TASENTRVMOHOVQVSNLQORSIPISLVLPV 959
 DB 912 QLELPVYAVYVMTSHGVSTKYLNF-TASENTRVMOHOVQVSNLQORSIPISLVLPV 971
 QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQC 1019
 DB 972 VLLNGVAVVDMVMEAPSQL--PCVSRKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029
 QY 1020 PFGIQEFPNATLKGNSLFDWYIKTSHNLLIYSTAEILFNDVSFTLLPGQAFVRSQTE 1079
 DB 1030 PFSVQBELDFTLKGNSLFGVRETLQKKVLSVSAEITFTSVYSQLPQGEAFPMQME 1089
 QY 1080 TKVEPPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFKKRYKDMSE 1128
 DB 1090 MVLEDEVYNAIPIMGSSVGALLLLALITATLYKLGFKKRYKEMLED 1138
 RESULT 15
 AAW65089
 ID AAW65089 standard; Protein; 1161 AA.
 XX AAW65089;
 AC AAW65089;
 XX 28-SEP-1998 (first entry)
 DT Human Beta-integrin alpha-d subunit protein isolated from clone 19A2.
 DE Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome;
 KW rheumatoid arthritis.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Peptide 1..16
 FT

```

FT Protein /label= signal
FT 17..1161 /note= "Beta-integrin alpha-D subunit"
FT Domain 17..1108 /label= extracellular
FT Domain 150..352 /note= "homologous to insertion domain of CD11 subunits"
FT Domain 1129..1161 /label= cytoplasmic
FT XX
PN US5728533-A.
XX
PN 17-MAR-1998.
XX
PD
XX
PF 07-JUN-1995; 95US-0485618.
XX
PR 07-JUN-1995; 95US-0485618.
PR 23-DEC-1993; 93US-0173497.
PR 05-AUG-1994; 94US-0286889.
PR 21-DEC-1994; 94US-0362652.
XX
PA (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Van DER VIEREN M;
PI
XX
DR WPI; 1998-206565/18.
DR N-PSDB; AAV35236.
XX
XX Screening assay for modulators of integrin binding - using
PT immobilised or labelled alpha-d polypeptide, useful for, e.g.
PT treating type-I diabetes
XX
XX Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a novel human beta-integrin alpha-d subunit.
CC This sequence is used in a method for identifying compounds that modulate
CC the interaction of alpha-d with a binding partner of alpha-d which
CC involves contacting an alpha-d polypeptide with an alpha-d binding
CC partner, one of which is immobilised and the other of which is labelled,
CC in the presence of a test compound, and determining if the compound
CC affects binding between the alpha-d polypeptide and alpha-d binding
CC partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress
CC syndrome and rheumatoid arthritis.
XX
SQ Sequence 1161 AA;
Query Match 57.8%; Score 3401; DB 19; Length 1161;
Best Local Similarity 59.3%; Pred. No. 3.5e-275;
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

QY 1 FNLDENAMTQENARGQSVQVQGSRRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 17 FNLDVEEPTIFQEDAGGQSVQVQGSRLVVGAPLEVVAAQNGRLYDCAAAQGMQPI 76
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHTQCSNTYVKGCLCFPGSNLRQQPQK 120
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 77 PLHIRPEAVNMSLGLTLAASNGSRLACGPTLHRVCGENSYSKGCILLGSRW-ETIQT 135
QY 121 FPEALRGCGQSDSIAPLIDGSGSII PHDFRMKEWSTVMEQLKSKTSLSMQYSEEP 180
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 136 VPDATPECPHOEMDIFVLIDGSGSIDQDNDFQMGKGFVQAVNGQPEGDTTLFALMOYNLL 195
QY 181 RIHFTKEFONPNPNSRLIKPTQLLGRTHATGLRKVVRLEFNITNGARKNAFKILFL 240
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 196 KIHFTFTQRTSPSQSLVDPILVKGTLTATGLTIVTQLFHHKNGARKSAKKILIVI 255
QY 241 TDGEFGDPLGVEDYIPELDREGVIRYVIGVDGAFRSEKSRQELNTVASKPPRHVFOIN 300
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 256 TDGQYKDPLEYSDVIPAERAGIIRYAIQVGHAFQGTARQELNTISSAPPDQHVFKVD 315

```

Search completed: November 25, 2003, 14:16:30
Job time : 40.2978 secs

```

QY 301 NPEALKTIQNLREKIFAIEGTQTGSSSFFHEMSQEGSAAITNSNGPLLTGVSYDWAG 360
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 316 NFAALGSIQKQLOEKIYAVEGTQSRASSSFFHEMSQEGSFTALTMDGLFLGAVGSFWSG 375
QY 361 GVFLVTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVANFR 420
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 376 GAFLYPPNMSPTFINMSQENVDMRDSYLGSTELAWKGVQNLVLGAPRYOHTGRAVFT 435
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYYQTRGGQVSVCP 480
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSNGSTDLILGAPHYYQTRGGQVSVCP 495
QY 481 PRGQARMQCDVAVLYGEOQPMGRFGAALTVLVDYNGDKLTDVAIGAPCEEDNRGAVYLF 540
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 496 PRGQVQVQCCDAVLRGEQGHGPMGRFGAALTVLGDVNEBKLDVAIGAPCEQENRGAVYLF 555
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTVGAQGHVLLLRSQ 600
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 556 HGASESGISPSHSQRIASSQLSPRIQYFGQALSSGQDLTQDGLMDLAVGARGOVLRLRSL 615
QY 601 PVLRYKAIMENPREVARNVFECDQVVKGEAGVRVCLHVQKSTRDRREGQIQSVVT 660
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 616 PVLKGVAMRFPVEVAKAVYRCWEEKPSALBAGDATVCLTIQKSSLDQL--GDQSSVR 673
QY 661 YDLALDSGRPHSRVFNENETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 674 FDLALDPGRLTSAIFNETKNPTLTRRTGLGHCETILKLLPDCVEDVWSPILHLNF 733
QY 721 SLVGTPLSAFGLNRPVLAEDAORLTALPPFKNGCNDNICODDLSITFSFMSLCLVVG 780
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 734 SLVREPIPSQNLRLPVLAVSQDLFTASLPFFKNCQDGLCEGDLGVTLSFSLQTLTVG 793
QY 781 GRPEENVTVVRNDGEDSVRTQVTFPPFLDLSVRKVTIQNORSORSWRLACESASSTEV 840
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 794 SSLELVIVTVWVWAGDSYGTVVSLYPAGLSHRRVSGAKQPHQSALALACETV-PTED 852
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKLLKAMVTSENNMPRTNKT 900
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 853 EG-LRSSRCSVNHPIFHEGSGNGTFTVTFDVSYKATGLGDRMLMRASSENKASSSKATF 911
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQSRSLPISLVFLVP 959
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 912 QLELPVKYAVYVMTISRQESTKYFNPATSDKKMKKEAHEHRYVNNLSQRDLAISINFWVP 971
QY 960 VRLNQTVIWDROPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCORTCDI 1019
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 972 VLLNGVAVMDVVMVMEAFSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQRCDV 1029
QY 1020 PFFGIQEEFNATLKGNSLDFDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQGGAFVRSQTE 1079
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 1030 PSFSVQEEELDFTLKGNLSFGWVRETLQKKVLVSVVAEITFTSVYSQLPQGEAFMRAQME 1089
QY 1080 TKVEPEVNPPLPIVGVSSVGGLLILALITAALYKLGFFPKRYQKDW MSE 1128
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 1090 MVLEEDVYNAIPIIMGSSVGALLLALITATLYKLGFKKRYKEMLED 1138

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:17:44 ; Search time 12.9635 Seconds
(without alignments)
4501.575 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884
Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKROYKDMSEGGPFGABPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 271250 seqs, 51324744 residues

Total number of hits satisfying chosen parameters: 271250

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

pred, No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5821.5	98.9	1152	5	US-09-592-617C-43
2	3444.5	58.5	1200	1	PCT-US03-28227-4834
3	3401	57.8	1161	7	US-60-517-843-2
4	3336.5	56.7	1108	1	PCT-US03-28227-4833
5	3308	56.2	1145	1	PCT-US03-28227-4832
6	3050.5	51.8	1047	7	US-60-487-610-1574
7	1546.5	26.3	1170	7	US-60-487-610-1880
8	1543.5	26.2	1170	5	US-09-592-617C-42
9	1166	19.8	413	7	US-60-487-610-1788
10	1144	19.4	1179	7	US-60-487-610-2494
11	1140	19.4	1179	1	PCT-US03-18234-2
12	1103.5	18.8	1149	7	US-60-487-610-1507
13	1093.5	18.6	1179	6	US-10-474-794-250
14	1055	17.9	1180	6	US-10-474-794-307
15	1054	17.9	1179	7	US-60-493-369-8
16	1054	17.9	1179	7	US-60-512-690-8
17	1054	17.9	1181	7	US-60-493-369-10
18	1054	17.9	1181	7	US-60-512-690-10
19	1052	17.9	1181	7	US-60-487-610-2692
20	1052	17.9	1181	7	US-60-485-450-1726
21	1052	17.9	1181	7	US-60-493-369-9
22	1052	17.9	1181	7	US-60-512-690-9
23	1051.5	17.9	1167	7	US-60-490-890-559
24	1050	17.8	1177	6	US-10-461-862-146
25	1048	17.8	1179	6	US-10-461-862-148
26	1043	17.7	1147	1	PCT-US03-00252A-42

27 1043 17.7 1147 6 US-10-336-603A-42 Sequence 42, Appl
28 1040 17.7 1147 1 PCT-US03-28227-4195 Sequence 4195, Ap
29 1035.5 17.6 1167 6 US-10-670-186-435 Sequence 435, Ap
30 855.5 14.5 987 1 PCT-US03-26780-1490 Sequence 1490, Ap
31 712.5 12.1 695 6 US-10-461-862-143 Sequence 143, Ap
32 650 11.0 1029 6 US-10-417-375-6 Sequence 6, Appl
33 650 11.0 1029 6 US-10-417-375-8 Sequence 8, Appl
34 643 10.9 1038 1 PCT-US03-18234-36 Sequence 36, Appl
35 637 10.8 1021 6 US-10-417-375-3 Sequence 3, Appl
36 636 10.8 1020 7 US-60-487-610-1825 Sequence 1825, Ap
37 634 10.8 1035 7 US-60-487-610-1603 Sequence 1603, Ap
38 627 10.7 1049 6 US-10-322-281-343 Sequence 343, Ap
39 599 10.2 1008 6 US-10-322-281-346 Sequence 346, Ap
40 537.5 9.1 1130 7 US-60-487-610-1902 Sequence 1902, Ap
41 537.5 9.1 1130 7 US-60-502-656-214 Sequence 214, Ap
42 537.5 9.1 1130 7 US-60-512-690-413 Sequence 413, Ap
43 537 9.1 1073 7 US-60-487-610-1903 Sequence 1903, Ap
44 537 9.1 1073 7 US-60-502-656-212 Sequence 212, Ap
45 537 9.1 1073 7 US-60-512-690-412 Sequence 412, Ap

ALIGNMENTS

RESULT 1
US-09-592-617C-43
; Sequence 43, Application US/09592617C
; GENERAL INFORMATION:
; APPLICANT: Amnaut, M. Amin
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN ANTAGONISTS
; FILE REFERENCE: 00786-267002
; CURRENT APPLICATION NUMBER: US/09/592,617C
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 08/380,167
; PRIOR FILING DATE: 1999-01-30
; PRIOR APPLICATION NUMBER: US 08/216,081
; PRIOR FILING DATE: 1994-03-21
; PRIOR APPLICATION NUMBER: US 07/637,830
; PRIOR FILING DATE: 1991-01-04
; PRIOR APPLICATION NUMBER: US 07/539,842
; PRIOR FILING DATE: 1990-06-18
; PRIOR APPLICATION NUMBER: US 07/212,573
; PRIOR FILING DATE: 1988-06-28
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -16 to -1
US-09-592-617C-43

Query Match 98.9%; Score 5821.5; DB 5; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSVVVVGAPQEIIVANQKSLYQCDYSTGSCBPI 60
|||||
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSVVVVGAPQEIIVANQKSLYQCDYSTGSCBPI 76
|||||
QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVKGICFLFGSNLRQOPK 120
|||||
DB 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVKGICFLFGSNLRQOPK 136
|||||
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWVSTVMBQLKKSTLFLSLMOYSEBF 180
|||||
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMBQLKKSTLFLSLMOYSEBF 196
|||||
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHRTATGURKVVVRELFTNGARKNAFKILFLL 240
|||||
DB 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHRTATGIRKVVVRELFTNGARKNAFKILVVI 256
|||||

Qy 241 TDGKFGDPLGYEDVPELDRGVIRVIGVGDFAFRSEKSRQELNVTASKPPRDHVFQIN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRVIGVGDFAFRSEKSRQELNVTASKPPRDHVFQIN 316
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEQTREGGQVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEQTREGGQVSCPL 496
Qy 481 PRGORARWQCDVILYGEQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEBEDNRGAYVLF 540
Db 497 PRG-RARWQCDVILYGEQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEBEDNRGAYVLF 555
Qy 541 HGTSGSGTSPSHSQRISAGSKSLPRLOYFGQSLSGQDITMDGLVDLTVGAGHVILLRSQ 600
Db 556 HGTSGSGTSPSHSQRISAGSKSLPRLOYFGQSLSGQDITMDGLVDLTVGAGHVILLRSQ 615
Qy 601 PVLRVKATMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIOQSVVT 660
Db 616 PVLRVKATMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIOQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 735
Qy 721 SLVGTPLSAFNGLRPLVAEDAQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 780
Db 736 SLVGTPLSAFNGLRPLVAEDAQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 795
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915
Qy 901 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHOYQVSNLQBSLPISLVFLVPV 960
Db 916 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHOYQVSNLQBSLPISLVFLVPV 975
Qy 961 RLNOTVIWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1020
Db 976 RLNOTVIWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1035
Qy 1021 FFGIOEEFNATLKGNSLSEFDWIKTSHNHLILVSTAEILFNDSVFTLLPGQAFVRSQDET 1080
Db 1036 FFGIOEEFNATLKGNSLSEFDWIKTSHNHLILVSTAEILFNDSVFTLLPGQAFVRSQDET 1095
Qy 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKROYKQMMSEGGPPGAEPO 1137
Db 1096 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKROYKQMMSEGGPPGAEPO 1152

RESULT 2

PCT-US03-28227-4834

; Sequence 4834, Application PC/TUS0328227

; GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION; SCHMIDT, Jeanette P.;

; APPLICANT: WRIGHT, Rachel J.; BRUNS, Christopher M.;

; APPLICANT: MARJANOVIC, Mirjana M.; SHEN, Fan.;

; APPLICANT: HARTSHORNE, Joanne A.; SUCHOROLSKI, Martin;

; APPLICANT: ALTUS, Christina M.; PITTS, Steven J.;

; APPLICANT: ELDER, Linda V.; MOONEY, Elizabeth M.;

; APPLICANT: DELEGEANE, Angelo M.; PANESAR, Iqbal S.;

; APPLICANT: BANVILLE, Steven C.; REDDY, Thirupathi P.;

; APPLICANT: STEVENS, Kristian A.; BLANCHARD, John L.;

; APPLICANT: PANZER, Scott R.; WANG, Xinhao;

; APPLICANT: AU, Alan P.; GERSTIN, Edward H., Jr.;

; APPLICANT: PERALTA, Careyna H.; ANDERSON, Scott E.;

; APPLICANT: RIOUX, Pierre; SHEN, Edward J.;

; APPLICANT: WU, Mingham C.; STUVE, Laura L.;

; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;

; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;

; APPLICANT: VITT, Ursula A.; KIRTON, Edward;

; APPLICANT: XU, Yuming; KWONG, Mary;

; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;

; APPLICANT: MA, Yan; JACKSON, Jennifer L.;

; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;

; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.;

; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: PN-0100 PCT

; CURRENT APPLICATION NUMBER: PCT/US03/28227

; CURRENT FILING DATE: 2003-09-12

; PRIOR APPLICATION NUMBER: US 60/410,260

; PRIOR FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: US 60/410,259

; PRIOR FILING DATE: 2002-09-12

; NUMBER OF SEQ ID NOS: 5444

; SOFTWARE: PERL Program

; SEQ ID NO 4834

; LENGTH: 1200

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: 965829.PT40p

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1) ... (1200)

; OTHER INFORMATION: unknown or other

; PCT-US03-28227-4834

Query Match

58.5%; Score 3444.5; DB 1; Length 1200;

Query Local Similarity 59.1%; Pred. No. 2.4e-191;

Matches 689; Conservative 142; Mismatches 292; Indels 43; Gaps 5;

Qy 1 FNLDTENAMTFQENARFGQSVVQOGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60

Db 20 FNLDTEBLTAFRVDSAGFGDSVVQYANSWVVGAPQKITAAQNTGGYQCGYSTGACPEI 79

Qy 61 RLQVPEAVNMVSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFLGSLNRQOPOK 120

Db 80 GLQVPEAVNMVSLGLSLAATTSPPOLLACGPTVHHCGRNMYLTGCLFLLGPT--QUTQR 137

Qy 121 FPALRGCPQEDSDIAFLIDGSGSIIPHDPRMKKEWSTVMEOLKKSCTLFSLMQXSEEF 180

Db 138 LPVSRQECPRQEQDIIVFLIDGSGSISRNPATMNFVRAVISOFORPSTQFSLMQFSNKF 197

Qy 181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVRELNITNGARKNAFKILFL 240

Db 198 QTHFTPEFRSSNPLSLASVHQLQGTFTATAIQNVVHRLPHASVAGARRDAKILIVI 257

Qy 241 TDEKFGDPLGYEDVPELDRGVIRVIGVGDFAFRSEKSRQELNVTASKPPRDHVFQIN 300

Db 258 TDGKFGDSDYKDVIPMAAAGIIRYAIQVGLAFQNRNSWKELNDIAKPSQEHIPKVE 317

Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360

Db 318 DFDKALQIQOLKEKIFAIEGTQTTSSSSFELEMAQEGFSAVFTPDGVPVGVGSFTWSG 377

Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420

Db 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGQSLVGLGAPRYOHTKAVIFT 437

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEQTREGGQVSCPL 480

Db 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDSNGSTDLVLGAPHYYEQTREGGQVSCPL 497

QY 481 PRQARWOCDAVLYGEOQPMGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAAYLIF 540
DB 498 PRGWR-RWNCDAVLYGEOQHMPGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAAYLIF 556
QY 541 HGTSGSGISPSHSORITAGSKLSPRLOYFGOSLSGGODLTMDGLVDLTGCAOQHVLILRSQ 600
DB 557 HGVLPSPISPSHSORITAGSKLSPRLOYFGOSLSGGODLTMDGLVDLTGCAOQHVLILRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFCNDQWVKGKEAGEVRVCLHVQKSTRDLRLEGQIQSVVT 660
DB 617 PVLWGVGSWQFPAEIPRSAFECEQVSEQTLVQSNICLYIDKRSKNLLSRDLQSSVT 676
QY 661 YDLALDSGRPHRAVENETKSTRQTQVLGTQTCETIKLQLPNCIEDPVPVILRLNF 720
DB 677 LDALDPGRLSPRATFOETKNRSLSRVVLGLKAHCENFNLLPSCEVDSVPTILRLNF 736
QY 721 SLVGTPLSAFNLRLPVLAEADQRLFTALPFPEKNCNDNICODDLSITFSFMSLCLVVG 780
DB 737 TLVGKELLAFNLRLPVLAEADQRYFTASLPFEKNCADHICQDNLGISFSFPGKSLVVG 796
QY 781 GPREFNVTVVRNDGEDSVYTOVTFEPLDLSYRKVSTLQNRORSRWLACESASSTEV 840
DB 797 SNLELNAEVMVNDGEDSVYTOVTFEPLDLSYRKVSTLQNRORSRWLACESASSTEV 854
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANYTSENMPRTNKTFF 900
DB 855 SQGTWSTSCRINHPIFRGQAOTFLATFDVSPKAVLGRLLLTANVSSNNTPRTSKTFF 914
QY 901 QLELPVKYAVYVWVTSHGVSSTKYLNTAS-ENTSRVNHQY-----940
DB 915 QLELPVKYAVYVWVTSHGVSSTKYLNTAS-ENTSRVNHQY-----940
QY 941 -----QVSNLQORSPLSLVLPVRLNQTIVWDRPQVTFSENLSST 982
DB 975 TGEFGVLWGLQCQPLXQVNNLQORDLPVSINFVWVPELVNQAQVMDVEVSHPPQPSLR 1034
QY 983 CHTKERLPSHSDPLAELRKAAPVNCIAVCORIQCDDIPFGIQEENFATLKNLSPDWYI 1042
DB 1035 CSSEKIAPASDPLAHIQKNPVLDCSIAGCLFRCDVPSFSVQEEELDTLKNLSPGWVR 1094
QY 1043 KTSNHLILVSTRAIILFNDSVFTLLPGQAFVRSQTEVKVEFEVNPPLIVGSSVGL 1102
DB 1095 QILQKRVSVSVVAEITFDTSVYSQLPQGEAFMRAQTTLVLEKYKHNPPLIVGSSIGGL 1154
QY 1103 LLLALITALYKLGFFKROYKDMSE 1128
DB 1155 LLLALITALYKVGFFRQYKEMME 1180

RESULT 3

US-60-517-843-2
; Sequence 2, Application US/60517843
; GENERAL INFORMATION:
; APPLICANT: Weaver, Lynne
; TITLE OF INVENTION: METHODS OF TREATING CHRONIC PAIN USING COMPOSITIONS THAT
; FILE OF INVENTION: SPECIFICALLY BIND CD11D (ALPHA-D) INTEGRIN
; FILE REFERENCE: 27866/35702
; CURRENT APPLICATION NUMBER: US/60/517, 843
; CURRENT FILING DATE: 2003-11-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-517-843-2

Query Match 57.8%; Score 3401; DB 7; Length 1161;
Best Local Similarity 59.3%; Pred. No. 7.6e-189;
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNRGLSYQCDYSTGSCPEI 60

DB 17 FNLDVEEPTIFQEDAGGFGQSVVQFGGRLVVGAPLEVAANQTRLVDCAAATGMCQPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVVKGLCFGLFNSLRQQQPK 120
DB 77 PLHIRPEAVNMSLGLTAASTNGSRLAACGPTLHRVCGENSYSKSGCLLGSRW-EIQT 135
QY 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEWSTVMEQLKSKTILFSLMOYSSEF 180
DB 136 VPDATPECPHOEMDITVFLIDSGSIDQDNFQMGKGFQVAVMGQFEGTDTLTFALMOYSNLL 195
QY 181 RHFTFKFQNNPNSRLIKPITOLGRTHTATGLRKVVRELFTNGARKNAKILFL 240
DB 196 KIHFTTFRTPSPQOOSLVDPIVQKGTFTATGLTIVTQTFHKGARKSAKILIVI 255
QY 241 TDGEFGPLGVEDYIPELDREGVIRYVIGVGDAFRSEKSRQELMTVASKPRDRHVQIN 300
DB 256 TDGQYKDPLEYSQVIPAOKAGIIRYAIKVGHAFQGPARTARQELNTISSAPQDHVFKVD 315
QY 301 NPEALKTIQONLREKIFAIEGTQSSSFHHEMSQEGFSAITNSGPLLSTVSGYDNAG 360
DB 316 NPAALGSIQKQLEKIXAYEGTQSRASSSFHHEMSQEGFSTALTMDGLFLGAVGFSWSG 375
QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIIILNRVQSLVLAGRYOHIGLVAMPR 420
DB 376 GAFLYPPNMSPTFINMSQENVDMSYLGSTELALWKGQVQLVLAGRYOHTGRAVPT 435
QY 421 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDVLVIGAPHYEQTRCGQVSCPL 480
DB 436 QVSRQWRKKAETVGTQIGSYFCASLCSVDVDSNGSTDVLVIGAPHYEQTRCGQVSCPL 495
QY 481 PRQARWOCDAVLYGEOQPMGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAAYLIF 540
DB 496 PRQARWOCDAVLYGEOQPMGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAAYLIF 555
QY 541 HGTSGSGISPSHSORITAGSKLSPRLOYFGOSLSGGODLTMDGLVDLTGCAOQHVLILRSQ 600
DB 556 HGTSGSGISPSHSORITAGSKLSPRLOYFGOSLSGGODLTMDGLVDLTGCAOQHVLILRSQ 615
QY 601 PVLVRKAIIMEFNPREVARNVFCNDQWVKGKEAGEVRVCLHVQKSTRDLRLEGQIQSVVT 660
DB 616 PVLVRKAIIMEFNPREVARNVFCNDQWVKGKEAGEVRVCLHVQKSTRDLRLEGQIQSVVT 673
QY 661 YDLALDSGRPHRAVENETKSTRQTQVLGTQTCETIKLQLPNCIEDPVPVILRLNF 720
DB 674 YDLALDPGRLSPRATFOETKNRSLSRVVLGLKAHCENFNLLPSCEVDSVPTILRLNF 733
QY 721 SLVGTPLSAFNLRLPVLAEADQRLFTALPFPEKNCNDNICODDLSITFSFMSLCLVVG 780
DB 734 SLVREPIPSQNLRLPVLAVGSQDLFTASLPFEKNCQDGLCEGLVTLVLSGLQTLTVG 793
QY 781 GPREFNVTVVRNDGEDSVYTOVTFEPLDLSYRKVSTLQNRORSRWLACESASSTEV 840
DB 794 SSLELNVTVVRNDGEDSVYTOVTFEPLDLSYRKVSTLQNRORSRWLACESASSTEV 852
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANYTSENMPRTNKTFF 900
DB 853 EG-LASSRCSVNHPIFHGSGNCTFTVTDVSKATLGDRLMARASSENKASSKATF 911
QY 901 QLELPVKYAVYVWVTSHGVSSTKYLNF-TASENTSRVNHQYQVSNLQORSPLISLPLVP 959
DB 912 QLELPVKYAVYVWVTSHGVSSTKYLNF-TASENTSRVNHQYQVSNLQORSPLISLPLVP 971
QY 960 VRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDPLAELRKAAPVNCIAVCORIQCDDI 1019
DB 972 VLLNGVAVMDVWEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQRCDV 1029
QY 1020 PFFGIQEEFNATLKNLSPDWYIKTSHNHLILVSTRAIILFNDSVFTLLPGQCAFVRSQTE 1079
DB 1030 PSFSVQEEELDTLKNLSPGWVRETLLQKVLVSVVAEITFDTSVYSQLPQGEAFMRAQME 1089
QY 1080 TKVEFEVNPPLIVGSSVGGLLLALITALYKLGFFKROYKDMSE 1128


```
; APPLICANT: WU, Mingham C.; STUVE, Laura L.;
; APPLICANT: LAGACE, Robert E.; SPIRO, Peter A.;
; APPLICANT: STEWART, Elizabeth A.; WINGROVE, James A.;
; APPLICANT: VITT, Ursula A.; KIRTON, Edward;
; APPLICANT: XU, Yuning; KWONG, Mary;
; APPLICANT: POLICKY, Jennifer L.; HURWITZ, Bonnie L.;
; APPLICANT: MA, Yan; JACKSON, Jennifer L.;
; APPLICANT: GIETZEN, Darryl; PATURY, Srikanth;
; APPLICANT: SHI, Xiaobing; SUAREZ, Charlyn J.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PN-0100 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/28227
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/410,260
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/410,259
; PRIOR FILING DATE: 2002-09-12
; NUMBER OF SEQ ID NOS: 5444
; SOFTWARE: PERL Program
; SEQ ID NO: 4832
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 965829.PT34P
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (1145)
; OTHER INFORMATION: unknown or other
PCT-US03-28227-4832

Query Match 56.2%; Score 3308; DB 1; Length 1145;
Best Local Similarity 57.5%; Pred. No. 1.8e-183;
Matches 670; Conservative 128; Mismatches 270; Indels 98; Gaps 6;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGRVVVVGAPQEIIVANQROSLYOCDYSTGSCPEI 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
20 FNLDEELTAFRVDSAGFGDSVVQVANSVVVVGAPQKITAANQTLGGLYQCGYSTGACEPI 79
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLGSLNLRQPOQK 120
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
80 GLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLGSLNLRQPOQK 137
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
121 FPEALRGCPQEDSDIAFLDGGSGIIPHDFRMKEWSTVMEQLKSKTFLSLMOYSBEF 180
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
138 LPVSQCEQPCQDQIVFLDGGSGISSRNFAFMNFVRAVISQFORPSTQSLMOPSKNF 197
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
181 RIHFTKQFQNNPNSLIKPIITQLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
198 QHFTFEFRSSNPLSLASVHQLQGGTYTATQNVVHRLFHASYGARRDAAKILIVI 257
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
241 TDGEKGDPLGYEDVPELDREGVTRYVIGUGDAFRSEKSRQELNTVASKPPRDHVQIN 300
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
258 TDGKEGSLDYKDVIPADAAIIRYALGVGLAFQNRNSWKELNDIAKPSQERHFKVE 317
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
301 NFEALKTIONQLREKIFALEGTQGTSSSFEHMSQEGFSAATISNGPLLSVTGVDWAG 360
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
318 DFDAKLDQNKQKEKIFALEGTQGTSSSFEHMSQEGFSAATISNGPLLSVTGVDWAG 377
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
361 GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
378 GAFLYPPNMSPTFINMSQENVDMDRSDYLSGYSFELALWKGVQSLVGLGAPRYQHTGRAVFT 437
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
421 QNTGWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAPHYYQTRGGQSVVCLP 480
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
438 QVSRQWRMAEAVTGIGSYFGASLCSDVDVDSNGSTDVLIGAPHYYQTRGGQSVVCLP 497
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
481 PRGQARMCCDAVLYGEGQPNWRFCAALTVLGDVNGDKLTDAVACAPCEENRCAVLYF 540
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
498 PRGWR-RMWCDAVLYGEGQPNWRFCAALTVLGDVNGDKLTDAVACAPCEENRCAVLYF 556
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
541 HGTSGSGISPSHRSQIRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLITVGAQHVLRLRSQ 600
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1574
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1574

Query Match 51.8%; Score 3050.5; DB 7; Length 1047;
Best Local Similarity 54.7%; Pred. No. 1.2e-169;
Matches 610; Conservative 146; Mismatches 256; Indels 103; Gaps 5;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGRVVVVGAPQEIIVANQROSLYOCDYSTGSCPEI 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
18 FNLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLGSLNLRQPOQK 120
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGCLFLGSLNLRQPOQK 120
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

; Sequence 1574, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1574
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1574
```


Db 78 PLHIREAVNNSLGLTLAABSTNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-BIIQT 136
Qy 121 FPEALRGCPEDSDIAFLIDSGSIIIPHPRRMKWVSTVMEQKKSKTFLFSIMQVSEEF 180
Db 137 VPATPECPHQEMDIIVFLIDSGSIDQNDFNQMKGFQVAMGQFEGTDTLFLAMQVSNLL 196
Qy 181 RIHFTPEFQNNPNRSLKPIITQLLGRHTATGLAKRVKRELFNITNGARKNAKFLIFLL 240
Db 197 KIHTFTQFTSQQSLVDPIVQLKGLTFTATGILTVVTQLPHKNGARKSAKKILIVI 256
Qy 241 TDEKFGDPLGYEDVPELDRGVIRVIGVDAFSEKSRQBELINTVASKPPRDHVFQIN 300
Db 257 TDQKQVPLEYSDVIPAQEKAGIIRYAIGVGHAFQCPARQELNITISSAPPDHVFQV 316
Qy 301 NFPAKTIQNLREKIFALEGTGTGSSSEPEHMSOEGFSAALTSNGPLLSVTGSDWDAG 360
Db 317 NFAALGSIQOLQEKIYAVGTQSRASSSFQHEMSOEGFSTALTMQGLFLGAVGFSWSG 376
Qy 361 GVFLYTSKESKSTFNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYOHLGLVAMFR 420
Db 377 GAFLYPNMSPFTINNSENVDNRDSYLGSTELAKWGVQNLVGLAPRYOHTGKAVIPT 436
Qy 421 QNTGMESNANVKTGTGAYFGASLCSVDVDSNGSTDVLVLIAPHYVETRGQVSVCP 480
Db 437 QVSRQWRKAEVGTGTQISYFGASLCSVDVDSNGSTDVLVLIAPHYVETRGQVSVCP 496
Qy 481 PRQBARWQCDVLYGEOGQPWGRFGAALTVLGDVNGDKLTVAIGAPGEDNRGAVYLF 540
Db 497 PRQORVQWQCDVLYGEOGQPWGRFGAALTVLGDVNGDKLTVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHRSORISPRLOYFGOSLGGQDLTMDGLVLTGVAQGHVLLRSQ 600
Db 557 HGASEGISPSHRSORISPRLOYFGOSLGGQDLTMDGLVLTGVAQGHVLLRSQ 616
Qy 601 PVLKRVKAIMBNPREVARNVECDQVVKGEAGEVRVCLHVQKSTRDRUREGOISVVT 660
Db 617 PVLKRVKAIMBNPREVARNVECDQVVKGEAGEVRVCLHVQKSTRDRUREGOISVVT 674
Qy 661 YDLALDGRPHSRVAVNETNSTRTQOVGLGTCTETKLOIPNCIEDPVSIVLRNF 720
Db 675 YDLALDGRPHSRVAVNETNSTRTQOVGLGTCTETKLOIPNCIEDPVSIVLRNF 734
Qy 721 SLVGTPLSAFQNLRPVLAEDAQRLLFTALFPFKXGNDNICODDLSTFMSLDCILVVG 780
Db 735 SLVREPIPSQNLRPVLAEGSODLFTASLPFKXGNDNICODDLSTFMSLDCILVVG 794
Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASTEV 840
Db 795 SSLELNVITVMNAGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASTEV 836
Qy 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTPE 900
Db 837 ----- 836
Qy 901 QLELPVKVAVVMVVTSHGVSTKYLNFTASENTSVMHOVOVSNLQORSLPISLVFLVP 959
Db 837 -----STKYNFNATSEKQKKEAHRVYNVNLQORSLPISLVFLVP 877
Qy 960 VRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDI 1019
Db 878 VLLNGVAVMDVWMEAPQSLL-PCVSEKPPQHSDFLTQISRSPLMDCSIADCLQFRCDV 935
Qy 1020 PFGIQEFNATLKNLSFDWYIKTSNHLIIVSTABILFNDVSTFLPQOGAFVRSQTE 1079
Db 936 PFSVQBELFTLKNLSFGVRETLOKVLWVSAEITDTSVYSQLPGQEAQFMAQME 995
Qy 1080 TKVEPVPNPLPLIVGSSVGLLLALITAAKYK 1114
Db 996 MVEEDEEVYNAIFIMSSVQALLLITATLYK 1030

RESULT 7

US-60-487-610-1880

; Sequence 1880, Application US/60487610

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,

; FILE REFERENCE: CL001469

; CURRENT APPLICATION NUMBER: US/60/487,610

; CURRENT FILING DATE: 2003-07-17

; NUMBER OF SEQ ID NOS: 97101

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1880

; LENGTH: 1170

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-487-610-1880

Query Match

26.3%; Score 1546.5; DB 7; Length 1170;

Best Local Similarity 34.3%; Pred. No. 9.4e-82;

Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;

Qy 1 FNLDTENAMTQ--ENARGFGQSVVOLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCSE 58

Db 26 YNLDVRGARSFSPRAGRHFGRVLRVQV-GNGVIVGAPGE---GNSGTGSLYQCQSGTGCHCL 81

Qy 59 PIRLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHOTCSENTYVVKGLCFGLFSGNLR--- 115

Db 82 PVTLR-GSNVTSKYLGMTLATPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGP 140

Qy 116 -QOQPKFPEALRGCPEDSDIAFLIDSGSIIIPHPRRMKWVSTVMEQKKSKTFLFSLM 174

Db 141 LQGRPGQFCIKG---NVDLVFLFDGNSLQDPDEFQKILDFMKVYKLSNTSYQFAAV 196

Qy 175 QVSEEPRIIHTKPEFQNNPNRSLKPIITQLLGRHTATGLAKRVKRELFNITNGARKNAF 234

Db 197 QFSTSYKTFEFDSDYVYKRDPPDALLKHVHMLLTNTFGAINTVATEVFEELGARPDAT 256

Qy 235 KILFLLTDEKFGDPLGYEDVPELDRGVIRVIGVDAFSEKSRQBELINTVASKPPRD 294

Db 257 KVLIIITDGE--ATDSGNIDAAKD-----IIRVIGIKHFKQTKESQETLHKFASRPASE 309

Qy 295 HVFQINNFEALKTQNLREKIFAIEGTQTGSSSEFHEMSQEGFSAALITSNGLLSTVG 354

Db 310 FVKILTFFELKXDLFTLQKIVYIEGTSKQDLTSFNMLSSSGISADLSRGHVAVGAVG 369

Qy 355 SYDNAGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQH 412

Db 370 AKDWAGGFLDLKADLQDDTFIGNPLTPEVRAGYLGTVTWLPSRQKTSLLASGAPRYQH 429

Qy 413 IGLVAMFR--QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLIAPHYVEOT 470

Db 430 MGRVLLFQBPQGGHWSQVQTHGTQISYFGELCGVDVDDGETELLIGALPLFYGEQ 489

Qy 471 RGGQSVCPPLPRQARWQCDAV--LYGEOGQPWGRFGAALTVLGDVNGDKLTDVAITGAP 528

Db 490 RGRVVIY-----QRRQLGFEVEVSELQDGPYPLGRFGEAITALTIDNGDLVDVAGAP 544

Qy 529 GEEDNRGAVLYFHGTSGSISPSHRSORISPRLOYFGOSLQSGQDLTMDGLVLT 588

Db 545 LEE--QGVVYIFNGRHG-GLSPQSQRIEGBTQVLSGQIOMFGRSHGVKDLGEGDLADAV 601

Qy 589 GAQCHVLLRSQPLVAKVKAIMEPNPREVARNVECDQVVK-CKEAGEVRCVCLHVQKSTR 647

Db 602 GAESQIMVLSRRPVDVMDVMTLMSFPAEIPVHEVECSYSTSNKMKEGVNTICFOI-KSLI 660

Qy 648 DRLREGOIOSVVTYDLALDSGRPHSRVAVNETNSTRTQOVGLGTCTETLKLQPNCI 707

Db 661 PQF-QGRVLVANLYTTLQDCHRTERRGLFPGRHRLRRNTAVT-TSMSCDTDFSFHFPVCV 718

Qy 708 EDPVSPIVLRNLNFSL---VGTPLS--AFGN-----LRPVLAEADAQRLLFTALFPFEKNCGN 757

GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1788
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1788

Query Match 19.4%; Score 1166; DB 7; Length 413;
Best Local Similarity 57.8%; Pred. No. 2.9e-60;
Matches 227; Conservative 61; Mismatches 103; Indels 2; Gaps 1;
Qy 1 FNLDTENAMTFQENARFGSGVQLOGSRVVGAPQEIIVAAANORGSLYQCXYSTGCEPI 60
Db 20 FNLDTTELTAFRVDSDAGSGSVQVYANSWVVGAPQKITAAQTGGYQCGYSTGACEPI 79
Qy 61 RLQVPEAVNMVSLGSLAAATTPPQLLACGPTVHQTCSNTYVKGCLFGLSGNLROQPK 120
Db 80 GLQVPEAVNMVSLGSLASTTSPQLLACGPTVHHECGRNWYLTGLCLLGLPT-QLTOR 137
Qy 121 FPEALRCPOEDSDIAFLIDGSGSIIPHDPRMKWSTVNEQLKSKTFLSLMQVSEEF 180
Db 138 LPVSRQBCPQEQDIFLIDGSGSISSRNFAFMNFVRAVISQFQRPSTQSLMQFSNKF 197
Qy 181 RIHFTPEFQNNPNRSLIKPITOLLGRTHATGLRVVRELFNITNGARKNAFKILFL 240
Db 198 QTHFTPEFRSSNPLSLASVHQLQGTTFATAIQNVHRLPHASTGAARDATKILIVI 257
Qy 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVGDFAFRSEKRGQELNVTASKPRDHVQIN 300
Db 258 TDGKKEGSLDYDVIPMAAAGIIRVAGVGLAFQNRNSWKLNDIASKPSQEHIPKVE 317
Qy 301 NFEALKTIQNLREKIFAEGTQTSSSSFEHEMSQSGFSAATNSGPLLSTVGSYDWAG 360
Db 318 DFDALKDIONLQKEKIFAEGTETSSSSFEHEMAQEGFSAVETPDGPVLGAVGFTWSG 377
Qy 361 GVFLYSKESKSTRNMTVRDSDMNDAYLGVA 393
Db 378 GAFLYPPNMSPTFINMSQENVDMDRSLGPPSA 410

RESULT 10
US-60-487-610-2494
; Sequence 2494, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2494
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-2494

Query Match 19.4%; Score 1144; DB 7; Length 1179;
Best Local Similarity 29.0%; Pred. No. 1.7e-58;
Matches 341; Conservative 213; Mismatches 451; Indels 172; Gaps 39;

Qy 45 GSLYQCDYSTGS--CEPI-RLQVP-----VBAVNMVSLGSLAAATTPPQLLACGPTVHQ 95
Db 65 GFLHRCSLVODEILCHPVEHVPIPKGRHGVTVVRSHHGVLICI-----QVLVRRP--HS 117
Qy 96 TCEYTYVYKGLCFGLSGNLROQPK-----SDIAFLIDGSGSIIPHDPRMKWSTVNEQL-- 119
Db 118 LSSELT--GTCSLILGPDRLRPOAQANFFDLENLLDPDARVDTGDCYSNKEGGEDDVNTA 174
Qy 120 KPPEALRGCPQED-----SDIAFLIDGSGSIIPHDPRMKWSTVNEQL-- 164
Db 175 RORRALEKEEEDKEEEDDEEBEAGTEIAIILDGSGSIDPPDFQRAKDFINMWRNFYE 234
Qy 165 KKSRTLFLSMQYSEEFRIHFTKFEQNNPNRSLIKPITOLLGRTHATGLRVVRELFN 224
Db 235 KCPECNFALVQYGVITQTEFDLQDSQDVMASTARVQNIITGVSVTKTASAMQHVLDSIFT 294
Qy 225 ITNGARKNAFKILFLITDGEKFGDPLGYEDVPELDRGVIRYVIGVGDFAFRSEKSRQL 284
Db 295 SSGSRKAKSKVMVVLTDGGIFEDPLNLTIVNSPMQGVVERFAIGVGEFKSARTAREL 354
Qy 285 NTVASKPPRDHVDQINNPEALKTIONOLREKIFAEGTQTSSSSFEHEMSQSGFSAAT 344
Db 355 NLIASDPDETHAFKVTNYMALDGLSLKRYNISWEGT---VGDALHYQLAQIGFSAQIL 411
Qy 345 SNGP-LLSTVSGSYDWAGGVFLY-TSKEKSTFINMTRVDSMDNDA-----YLGAAAIILRN 398
Db 412 DERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTAAAAAADAEMAAQYSLGVAVVLHKT 471
Qy 399 RVQSLVILGAPRYQHIGLVAMFR-QNTGMWESNANV-KGTQIGAYFCASLCSVDVDSNGST 456
Db 472 CSLSYVAGAPRYKHHG--AVFELQKEGREASFLVLEGEQMGSYFSGELCPVDIDMDGST 529
Qy 457 DLVLIGAPHYEQTRCGQVSVCPPLRGQARWQCDALVYGEQOPMGRFGAALTVLGDVN 516
Db 530 DFLVVAAPYHVHGEGRVYVYRLSE-QDGSFSLARILSGHPGFTWARFGFAMAANGDLS 588
Qy 517 GDKLTDVAITGAP---GEEDNR--GAVLYFHTGSGSISPSHSQRIAGSKLSPLRYFGQ 570
Db 589 QDKLTDVAITGAPLEGGFADGASFGSVIYNG-HWDGLSASPSQIRASTVAPGLQYFGM 647
Qy 571 SLSSGGQDLTMDGLVDLTGCAQGHVLLRSQPLRVKAIWEPNPREVARNVFECNDQVKG 630
Db 648 SMAGSFDISGDGLADITVGLQAVVFRSRPVVRLKVSMAFTPSALP-----IGF 697
Qy 631 KEAGEVRVCLHVQKSTRDLREGQIOSVVTYDLALDSGRPHSRSAVNETKNSTRRTQVL 690
Db 698 NGVVNVLCFEI-SSVTTASEGLREALNFTLDVDVGKRRRLQCSVRSCLGCLREWS 756
Qy 691 GLTQTCETLKLQLPN---CIEDPVSPVILRNFLNFSLVGTPLSAFAGNLRPVLAEDAQLFT 746
Db 757 SGSQLCEDL-LLMPTGELCEEDCFNSAVKVSQYL-QTPEGQTDHPQPLDRYTEPFAI 814
Qy 747 ALFPPEKNCNDNICODDLSITPFSNLDCLVVGPPREFNVTVTVNDGEDSRYTQVTF 806
Db 815 FQLPYEKACKNKLFCVAELQLA-TTVSQDELVVGLTKELTNLNTNSGEDSYMTSMALN 873
Qy 807 FPLDLSYRKVSTLQONRSQSRWRLACESASSTEVSGALKSTCSINHPIPPENSEVTNI 866
Db 874 YPRNLO-----LKRQKPPSPNIQCDDPQPV---ASVLMNCRIQHPVL-KRSSAHVS 923
Qy 867 TFDVDSKASLGNKLLKANVTSENN-----MPRTNKTFFQ---LELPVKYAVYVMTSHG 919
Db 924 VMQLEENAFNRTADITVTVNSNERWSLANETHTLQFRHGFVAVLSKPSIMYVNTGQGL 983
Qy 920 S---TKYLNFTASNTSRVMQHQVQNSLORSPLISLVFLVPVRLNQTIVWDRPQTFSE 977
Db 984 SHKBEFLFHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAVVKKLRTQ 1028
Qy 978 NLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIPFFGQEBFNATLKNLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQHVVEHWSVSCVIA-----SDKENVTVAEIS 1073

```
QY 1038 FDWIKTSHNHLIVST-----AEILFNDVSFTLLPGQAFVRSQETKVEPEVNPPL 1091
Db 1074 WD-----HSEELLKDVTEQLIGELISFNKSLYGLNAENH--RTKITVVFLLKDEKHS 1125
QY 1092 PLVIGSSVGGLLALITALAALYKLGFFKROYKOWMSE 1128
Db 1126 PIIKGSVGGLLVILVILFKCGFFRKRYOQLNLE 1162

RESULT 11
PCT-US03-18234-2
; Sequence 2, Application PC/TUS0318234
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: Antibodies that bind alphaE Integrin
; FILE REFERENCE: 1855-2025002
; CURRENT APPLICATION NUMBER: PCT/US03/18234
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 10/173,551
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(18)
PCT-US03-18234-2

Query Match 19.4%; Score 1140; DB 1; Length 1179;
Best Local Similarity 28.9%; Pred. No. 2.8e-58;
Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSYQCDSYSGS--CEPI-RLQVP-----VEAVNMSLGLSLAATSPPOLLAGCPVHQ 95
Db 65 GPLHRCSLVQDEILCHVEHPVPIKGRHGVTVVRSHHGLICI-----QVLVRP--HS 117
QY 96 TCSNTVVKGLCLFLGNSLHQPO-----119
Db 118 LSELT---GTCSLGPDLPQAQANFFDENLDDPARVDTGDCYNKEGGEDDVNTA 174
QY 120 KPEALRGCPQED-----SDIAFLDGGSGIIPHPFRMKWVSTWMEQL-- 164
Db 175 RQRALEKEEEDKEEEDDEEAEAGTEIAIILDGSGSIDPPQRAKOFISNMNRNFE 234
QY 165 KSKTLFSLMOYSEEFRIHFTFEFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFN 224
Db 235 KCFECNFAVQGVIGVQTEFDLRDSQDVMASLARVQNIQVGSVTKTASAMQHVLDISFT 294
QY 225 ITNGARKNARKIILLTDGKFGDPLOYEDVPELDREGVIRVIGVDGAFRSEKSRQEL 284
Db 295 SSHGSRKASKVMVLLTDGIFEDPLDITVINSPKMGYVERFAIGVGEFEKSAAREL 354
QY 285 NTVASKPRPDVQINNFEALKTIONOLREKIFAETGQTGSSSSPEHEMSQSGSAIT 344
Db 355 NLIASDPDETHAFKVTNMAIDGLLSKLRNIIISMEGT---VGDALHYQLAQIGFSAQIL 411
QY 345 SNGP-LISTVGSYDWMAGGVFLY-TSKEKSTFINNTRVDSIDMND-----YLGAAAAIILRN 398
Db 412 DERQVLLGAVGAFDWSGGLLYDTRSPRGRFLNQTAADAAAEAAQVSYLYGAVAVLHKT 471
QY 399 RVQSLVLGAPRYOHIGLVAMFR-QNTQWESNANV-KGTQIGAYFGASLCSDVDNSGST 456
Db 472 CSLSYVAGAPQYKHG--AVFELQKEGREASFLPVLEGEQMGSGYFSGSELCPVIDIDMDGST 529
QY 457 DLVLIGAPHYVEQTRGQVSVCPPLRGORARWOCDAVLVYCEQCPWGRFGAALTVLGDVN 516
Db 530 DFLVRAAPFTHVHGEGRVTYVRLSE--QDGSFSLARILSGHPGFTNARFGFAAAMGDLS 588
QY 517 GDKLTDVAIGAP-----GEEDNR--GAVYLPFGTSGSGISPSHSQRIAGSKLSRPLQYFGQ 570
```

```
Db 589 QDKLTDVAIGAPLEGFGADDGASFGSVIYNG-HMDGLSASPQRIRASTVAPGLQYFGM 647
QY 571 SLSSGGDLTMDGLVDLTGCAQGHVLLRSQPLVRKVAIMEFNPREFVARNVFECDQVVKG 630
Db 648 SMAGGFDISDGLADITVGTGQAVVFRSRPVVRLKVSMAFTPSALP-----IGP 697
QY 631 KEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDIALDSGRPHSRVAVNETKNSTRRTQVVL 690
Db 698 NGVVNVRLCFEI--SSVTTASESGLEALLNFTLDVVGKORRLQCSVRSCLGCLREWS 756
QY 691 GLTOTCETLKLQLPN-----CIEDPVSPIVLRNLSLVGTPLSAFAGNLRPVLAEDAORLFT 746
Db 757 SGSQLCEDL-LLMPTGEGELCEDCFSNASVKVSYQL-QTPEGQTDHPQILDRTYTFPAI 814
QY 747 ALFPPEKNGNDNICODLSITFSFMSLDCLVVGGRPFNFVTVTVNDGEDSVRTQVTF 806
Db 815 FQUPYKACKNKLFCVAELQLA-TTVSQQLVVGLTKELTNLNLNLSGEDSVMTSMALN 873
QY 807 FPLDLSYRKVSTLQNRORSWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNI 866
Db 874 YPRNLQ-----LKRQKPPSPNIQDDPPQV---ASVLIMCRIGHVPL-KRSSAHVSV 923
QY 867 TPDVDSKASLGNKLLKANVTSENN-----MPRTNKTEPQ---LELPVKYAVVMVVTSHGV 919
Db 924 VMOLEENAFNPRTADITVTVTNSNERRSLANETHTLQFRHGFVAVLKSPKSMYVNTQGL 983
QY 920 S--TKYLNFTASENTRVMQHOYQVSNLQGRSLPISLVLVPLVRLNQTWINDRPOVTFSE 977
Db 984 SHHKEFLHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAAVKLRTQ 1028
QY 978 NLSSTCHTKERLPSHSDFLAELRKAPVNCSTIAVCQRIQCDIPFFGIGEEFNATLKNLS 1037
Db 1029 ASIVCTWSQERACAYSS-VQHVVEHWSVCVTA-----SDKENVTVAAEIS 1073
QY 1038 FDWIKTSHNHLIVST-----AEILFNDVSFTLLPGQAFVRSQETKVEPEVNPPL 1091
Db 1074 WD-----HSEELLKDVTEQLIGELISFNKSLYGLNAENH--RTKITVVFLLKDEKHS 1125
QY 1092 PLVIGSSVGGLLALITALAALYKLGFFKROYKOWMSE 1128
Db 1126 PIIKGSVGGLLVILVILFKCGFFRKRYOQLNLE 1162

RESULT 12
US-60-487-610-1507
; Sequence 1507, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Honglin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1507
; LENGTH: 1149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1507

Query Match 18.8%; Score 1103.5; DB 7; Length 1149;
Best Local Similarity 27.7%; Pred. No. 3.5e-56;
Matches 343; Conservative 212; Mismatches 485; Indels 197; Gaps 44;

QY 1 ENLDTENAMTFQENARG-EGQSVVOL---QGSRVVVGAPQEIIVAAANGRSGLYCDYSTGS 56
Db 1 FNVDVKNVMTSGFVEDMFGYTVQQYENEEGKWLIGSLVPGPKRKTGDVYKPCVGRGE 60
QY 57 CEP-IRLQVPVEA-----VNMSLGLSLAATSPPOLLAGCPVTHQTCSENTVYVKG 106
```

Db 61 SLPCVKLDLPVNTSIPNTEVKENMTFGSTL-VTNPNGGFLACGPLYAYRCGLHYTTGI 119
Qy 107 CFLFGSNLRQOPKPEALRCPOEDSDIAFLIDGSGSIIIPHDFRMKEWSTVMEQLK- 165
Db 120 CSDVSPTEFQVNSIAP--VQECSTQ-LDIIIVLDGNSIYPWD--SVTAFNLNLERMDI 174
Qy 166 -KSKTLFSLMOYSEBFRIHFTKFPQNNPNRSLIKPITQLLGR--THATGLRKVVRLEF 223
Db 175 GPKOTQGVQYGENVTHERNLNKSSTEELVAAKKIVQGGQTMTALGIDTARKEAF 234
Qy 224 NITGARKNAKPIKLLTDGSKPGDPLGYEDVPELDREGVIRVIGVGDAPR-----SE 278
Db 235 TEARGARRGVKVMVIVTDGESH-DNHLKKVIOCDENIQRFSAIILGSRNGLSTE 293
Qy 279 KSRQELNVTASKPRDRHVQINNFEALATIQNLREKIFAIEGTQSGSSSFEHEMSQEG 338
Db 294 KPEVEIKSIASEPTKHFNFVSDALVTVIKTLGERIFALEATADQSAASFEMMSQTG 353
Qy 339 FSAALITNGPLISTVSGSDWAGGVLYTSKE-----KSTF-INMTRVDSMDNDAYLGAA 392
Db 354 FSAHYSQDWMVGAVDNGTVMVQKASQIIIPRNTTFNVSTKNEPL-ASVLYGTV 412
Qy 393 AILLRNVQSL-VLGAPRYOHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSDVD 451
Db 413 NSATASSGDVLYIAGQPRYNTGTQVVIYRMEDGNIKILQTLGSEQIGSYFGSILTTDID 472
Qy 452 SNGSTDVLILGAPHY-----YEQTR-GGOVSVCPPLPRQARWQCDVLY 495
Db 473 KDSNTDILLVAGAPMYGTEKEEQKVYVVALNQTREYOMSLPIKOTCCSSRQHNSTT 532
Qy 496 GEOGQPMW-RFGAALTVLVDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGISPSHSQ 554
Db 533 ENKNEPCGARGTAAVKDLNLDGNDIVIGAPLEDHGGAVIYHG-SGKTIRKEYAQ 591
Qy 555 RIAGSKLPRLOYFGOSLGGODLTWGLVDLTVGAQGHVLLRSQPLRVLRKAIEMFNPR 614
Db 592 RIPSQGGDKTKLFFQGSIHGMDLNGDGLTDTIGLGAALFWSRDAVAVKVTMNFEPN 651
Qy 615 EVARNVFECNQVVGKAG--EVRVCLHQ-KSTRDLRREGQIQSVTVYDLADSGRPH 671
Db 652 KVINIQKNCH--MEGKETVCINATVCFDVKLSKEDTIYEADLQ----YRTLDSLRQI 704
Qy 672 SRAVENET-----KNSTRROTQVLGLTQTCTELKLQLPNCIEDPVSPVLRNLFSLVGT 725
Db 705 SRSFSGTQERKQVNRITVRKE-----CTKHSFMYLHFDQDSVR---ITLDENLI-D 753
Qy 726 PLSAFENLRPVLAEDAORLFTALPFERKNCNDNICQDDLSITFSFMSLDCLVGGPRE- 784
Db 754 PENG-----PVLDDSLPNSVHEYIPFAKDCGNKEKICISDLJHVAATTEKDLIIVRQNDK 808
Qy 785 FNVTVVRNDGEDSYRTQVTFPPFLDLSVRKVSTLQNRQSRWSPLACESASSTEVSGAL 844
Db 809 FNVSLTVKNTKDSANTRTIYHVSPLNLFVSGIEAQKD-----SCSEN----- 851
Qy 845 KSTSCSINHPIIPENSEVTFTITFDVDSKASLGN-KLLKANVTSENMPRTNKTEFOLE 903
Db 852 HNICTKGVFPFLRGEMVTFKLFQNFYSYLMENVTVIYLSATSDSEEPETLSDNVNIS 911
Qy 904 LPKVAVVMVTVSHGVSTKYLNTASNTSRVMQHOYQVSN-----LGORS-----L 950
Db 912 IPVKEVGLQFYS-SASEYHISIAANETVPEINSTEIDGNEINIFYLIRKSGSPMPPEL 970
Qy 951 PLSLYF-----LVPVRLNQTWDRPQVTFSENLSSTCHTKE-----RLPS 991
Db 971 KLSISFPNMTSNGYVLPYPTGLSS-----SENANCRPHIFDPFINSKGMWT 1019
Qy 992 HSDFLAELRKAPVNCSTAVCORIOCDIPFFGQIE-----EFNATLK 1033
Db 1020 STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILWKTFFIKSYFSSNLIR 1076
Qy 1034 GNLSFDWYIKTSHNLLIVSTAEILFNDSVFTLLPQCGAFVRSQETKVPPEVFNPLPL 1093

Db 1077 GEL-----RSNASLVJSSN-----QKRELAIQISKOGLPORVPL 1112
Qy 1094 --IVGSSVGLLLALITAAALYKLGFFKROYKDMWSE 1128
Db 1113 WVILLSAFAGLLLLMLLILALWKIGFFKRLPKKKMEK 1149
RESULT 13
US-10-474-794-250
; Sequence 250, Application US/10474794
; GENERAL INFORMATION:
; APPLICANT: Carson-Walter, Eleanor
; APPLICANT: St. Croix, Brad
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107,00179
; CURRENT APPLICATION NUMBER: US/10/474,794
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/308,829
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 359
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-794-250

Query Match 18.6%; Score 1093.5; DB 6; Length 1179;
Best Local Similarity 27.8%; Pred. No. 1.4e-55;
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

Qy 1 FNLDTEVAMTFQENARG-FQGSVVQL---OGRSVVVCAPQEIIVANORGSLYQCDYSTGS 56
Db 29 FNVVDKNSMTFGSPVEDMFGYTVQYENEEGKWVLIGSLVGPQPKNRTGDKYKCPVGRGE 88
Qy 57 CEP-IRLQVPEA-----VNMSLGLSLAATTSPPQLLACQPTVHQTCSENTYVKGL 106
Db 89 SLPCVKLDLPVNTSIPNTEVKENMTFGSTL-VTNPNGGFLACGPLYAYRCGLHYTTGI 147
Qy 107 CFLFGSNLRQOPKPEALRCPOEDSDIAFLIDGSGSIIIPHDFRMKEWSTVMEQLK 166
Db 148 CSDVSPTEFQVNSIAP--VQECSTQ-LDIIIVLDGNSIYPWD-----VTAFLNLLK 198
Qy 167 -----SKTLFSLMOYSEBFRIHFTKFPQNNPNRSLIKPITQLLGR--THATGLRKV 219
Db 199 RMDIGPKQTQGVIGYGENVTHERNLNKSSTEELVAAKKIVQGGQTMTALGIDTAR 258
Qy 220 RELNITNGARKNAKPIKLLTDGCKFGDPLGYEDVPELDREGVIRVIGVGDAPR--- 276
Db 259 KEATFARGARGVKVMVIVTDGESH-DNHLKKVIOCDENIQRFSAIILGSRNGLSTE 317
Qy 277 --SEKSRQELNVTASKPRDRHVQINNFEALATIQNLREKIFAIEGTQSGSSSFEHEM 334
Db 318 LSTKFEVEIKSIASEPTKHFNFVSDALVTVIKTLGERIFALEATADQSAASFEMEM 377
Qy 335 SQEGFSAALITNGPLISTVSGSDWAGGVLYTSKE-----KSTF-INMTRVDSMDNDAYL 388
Db 378 SQTGFSAHYSQDWMVGAVDNGTVMVQKASQIIIPRNTTFNVSTKNEPL-ASYL 436
Qy 389 GYAAAIILNRVQSL-VLGAPRYOHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLC 447
Db 437 GTVNSATSSGDVLYIAGQPRYNTGTQVVIYRMEDGNIKILQTLGSEQIGSYFGSILTT 496
Qy 448 VDVSNGSTDVLILGAPHY-----YEQTR-GGOVSVCPPLPRQARWQCD 491
Db 497 TDIKDSNTDILLVAGAPMYGTEKEEQKVYVVALNQTREYOMSLPIKOTCCSSRQHN 556
Qy 492 AVLVEGEOQPMW-RFGAALTVLVDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGISP 550

Db 557 SCTTENKNEPCGARGFTAI AAVKDLNDLGDNDIVIGAPLEDDHGGAVYIHG-SGKTIRK 615
Qy 551 SHSQRIAGSKLSRLQYFGQSLSGGODLTMDGLVDLTGAGQGHVLLRLSRQVLRVKAIME 610
Db 616 EYAQRIPSGGDKTLKFFQSGIHGEMDNGDGLTDVTIGLGAALFWSRDAVAVKVTWN 675
Qy 611 FNPREVARNVFECDQVVKKEAG--EVRVCLHVQ-KSTRDRLRBQIQISQVVTYDLALS 667
Db 676 FEPNKVNIQKNCH--MECKETVCINATVCFEVLKSKEDTIYEADLQ----YRVTLDS 728
Qy 668 GRPHSRVAFNET-----KNSTRQTOVLGLTOTCETLKLQPNCEIDPSPVILNFS 721
Db 729 LROISRSFFSGTOERKQVRNITVRKSEC-----TKHSFYMLDKHDFQDSVR---ITLDFN 780
Qy 722 LVGTPLSAFAGNLPRVLAEDQRLFTALFPPEKNCNDNICQDDLSITTFSPMSDCLVVG 781
Db 781 LT-DPENG-----PVLDDSLPNSVHEIYIPAKDCGNKERCISDLSLVATTEKDLIVRS 834
Qy 782 PRE-FNVTVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLACSASSTEV 840
Db 835 QNDKFNVSILTVKNTKDSAYNTRTIVHYSPLNLFSGIEAIQKD-----SCSEN----- 881
Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGN-KLLKANVTSENMPRTNKTE 899
Db 882 ----HNITCKGVFFLRRGBMVTFKILFQNTSYLMENVTIYLSATSDSEEPETLSDNV 937
Qy 900 FOELPVPKYAVYVMTSHGVTSTYKLNFTASENTSRVMOHYOVSN-----LGQRS---- 949
Db 938 VNTSIPKVEVGLQFVS-SASEVHISIAANETVEVINSTEDIGNINIFYLIRKSGSP 996
Qy 950 ---LPISLVF-----LVPRLNQTVIMDRPQVTFSENLSSTCHTKE----- 997
Db 997 MPBLKLSISPNMTSNGYPVLYPTGLSS-----SENANCRPHIFEDPFSINSGK 1045
Qy 988 RLPSSHDFLAELKAPVWNGSIAVCQRIQCDIPFFGIE-----EPN 1029
Db 1046 KMTTSTD---HLKRGITLDCNTCKFATITCNLTSSDISQVNVSLILKWKPTFIKSYFSSLN 1102
Qy 1030 ATLKGNLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPQGAFAVRSQETKVPPEVFN 1089
Db 1103 LTRGEL-----RSENASLVSSN-----QKRELAIQISKDGLPG 1138
Qy 1090 PLPL--IVGSSVGLLLALITALYKLGFFKRYQKDMSE 1128
Db 1139 RVPLWILLSAFAGLLLLLMLLALWKGIFGFKPLKKOMK 1179

RESULT 14

US-10-474-794-307
; Sequence 307, Application US/10474794
; GENERAL INFORMATION:
; APPLICANT: Carson-Walter, Eleanor
; APPLICANT: St. Croix, Brad
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00179
; CURRENT APPLICATION NUMBER: US/10/474,794
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/308,829
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 359
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 307
; LENGTH: 1180
; TYPE: PRN
; ORGANISM: Rat
US-10-474-794-307

Query Match 17.9%; Score 1055; DB 6; Length 1180;
Best Local Similarity 27.4%; Pred. No. 2.3e-53;

Matches 346; Conservative 196; Mismatches 475; Indels 246; Gaps 48;
Qy 1 FNLDTENAMTFOENARG-FGQSVWLQ---QGSRVVVVGAPQEIIVAAANQRGSLYQCDSYSGS 56
Db 29 FNVVDVKNSSFGSPVEDMEGYTVQYENEEGKWLIGSPLVGPQPKARTGDVYKCPVGR 88
Qy 57 CSP-IRLVQVPEA-----VNSLGLSLAATSPOLLACGPTVHTQTCSENTYVKGL 106
Db 89 AMPCVKLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNNGGFLACGPLYAVRCGHLHYTTGI 147
Qy 107 CFLFSGNLRRQPKPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKEMVSTM---EQ 163
Db 148 CSDVSPTFOVNSFAP--VQECSTO-LDIVIVLDGNSIYP-----WESVIAFLNDL 196
Qy 164 LKK-----SKTLFSLMOYSEERIHFTKEFQNNPNRSLIKIPITOLLG-RHTATGLRK 217
Db 197 LKRMIDIGPKQTQGVIGYGENVTHEFNLNKYSTEEVLVAANKI GRQGLQMTALGIOT 256
Qy 218 VVRELFNITNGARKNAFKILFLTQGEKPGDPLGYEDVIPLEDRREGVIRYIVGVGDAFR- 276
Db 257 ARKEAFTARGARRGVKKVMVITDGESH-DNYRLKQVIQDCEDENIQRFSTAILGHYNR 315
Qy 277 ----SEKSRQELNTVASKPPRDHVFIQINNFEALKTIONLREKI PAIECTQTGSSSSFEH 332
Db 316 GNLSTEKFVEETKSTASEPTEKHFNVSDDELALVTIVKALGERIFALEATADQSAASPEM 375
Qy 333 EMSQGFSAATISNGPLLSLTVSGYDWAGGVFLYTSKEKSTFINMT--RVDSDND---AY 387
Db 376 EMSQGFSAHYQSDWMLGAVGAYDNQTVVMQKANQMVIPHNTTFTQTEPAKNEPLASY 435
Qy 388 LGYAAAILLRNVQSLVGLGAPRYQHIGLVAMFRONTGMMESNANVKGTOIGAYFGASLCS 447
Db 436 LGYTVNSATIPGDVLYIAGAPRYNHTGQVVIYKMGDNINILQTLGGEIGSYFGSVLTT 495
Qy 448 VDVDSNGSDLVILGAPHY-----YEOTR-GQOVSVCLPLRQORARWQCD 491
Db 496 IDIDKSDYTLGLLVGAPMYMTGTEKEEQKVYVYVYNQTRFYOMSLEPIRQTCSSSLKON 555
Qy 492 AVLYEEOGQPMG-RFGAALTVLGDVNGDKLTDVAIGAPCEEDNRGAVLYPHGTSGSGISP 550
Db 556 SCTKENKNEPCGARGFTAI AAVKDLNDVDFNDVIGAPLEDDHAGAVYIHG-SGKTIRE 614
Qy 551 SHSQRIAGSKLSRLQYFGQSLSGGODLTMDGLVDLTGAGQGHVLLRLSRQVLRVKAIME 610
Db 615 AVAQRIPSGGDKTLKFFQSGIHGEMDNGDGLTDVTIGLGAALFWARDVAVAVKVTWN 674
Qy 611 FNPREVARNVFECDQVVKKEAG--EVRVCLHVQ-KSTRDRLRBQIQISQVVTYDLALS 667
Db 675 FEPNKVNIQKNCH--VEGKETVCINATMCFHVKLSKEDSIYEADLQ----YRVTLDS 727
Qy 668 GRPHSRVAFNET-----KNSTRQTOVLGLTOTCETLKLQPNCI-----EDPVS 712
Db 728 LROISRSFFSGTOERKQVRNITVRSE-----CIRHSFYMLDKHDFQD 770
Qy 713 PVLRLNFSLVGTPLSAFAGNLPRVLAEDQRLFTALFPPEKNCNDNICQDDLSITTFSPM 772
Db 771 SVRVTLDFNLT-DPENG-----PVLDDALPNSVHEIYIPAKDCGNKERCISDLSLVNVT 824
Qy 773 SLDCLVVGGPRE-FNVTVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLA 831
Db 825 EKSLIVKSHQDKFNVSILTVKNTKDSAYNTRTIVHYSPLNLFSGIEAIQKD-----S 876
Qy 832 CESASSTSVSGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLL-LKANVTSEN 890
Db 877 CESN-----QNITCRVGFPLRAGETVTFKIIQFNTSHLSNAILHLSATSDSEE 927
Qy 891 NMPRTNKTEFOLELPVKYAV---YMWVTSHGVS-----KYLNTFASENTSRVMOHQ 939
Db 928 PLESNDNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNINVF 985
Qy 940 YQVSNLGRQSLP---ISLVF-----LVPRLNQTVIMDRPQVTFSENLSSTCHTKE----- 971
Db 986 YTIKRGHPMPPELQSLISFPNLDTADGPVLYPIG-----WSSSDNVNCRPSRLEDPFG 1039

```
Qy 972 -----QVTF-----ENLSTCHTKERLPSHSDFLAELKAPVWNCIAVCQRIQ 1016
Db 1040 INSGKWTISKSEVLKRGTIQDCSTC-----GVATITCSLLPSLSQ 1082
Qy 1017 CDI-----PFGIOEERF---NATLKNLSPDWIKTSHNHLLIVSTABILFNDVSFTLL 1067
Db 1083 VNVSLMLWKPTF-IRAHFSSNLTLRGELK-----SENSLTLSSN----- 1123
Qy 1068 PQGAFVRSQETKVEPEVENPLPL--IVGSSVGGGLLLALITAALYKLGFFKROYKDM 1125
Db 1124 -----RKRELAIOISKDGLPGRVPLWVILLISAFAGLLMLLILALWKIGFFKRPKKK 1177
Qy 1126 MSE 1128
Db 1178 MEK 1180

RESULT 15
US-60-493-369-8
; Sequence 8, Application US/60493369
; GENERAL INFORMATION:
; APPLICANT: Bruce DOMON
; APPLICANT: Tao HE
; APPLICANT: Xiaolong ZHANG
; APPLICANT: Karen KETCHUM
; APPLICANT: Ian McCAFFERY
; TITLE OF INVENTION: PANCREATIC DISEASES TARGETS AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001476
; CURRENT APPLICATION NUMBER: US/60/493,369
; CURRENT FILING DATE: 2003-08-08
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-493-369-8

Query Match 17.9%; Score 1054; DB 7; Length 1179;
Best Local Similarity 26.7%; Pred. No. 2.6e-53;
Matches 326; Conservative 219; Mismatches 494; Indels 182; Gaps 44;

Qy 1 FNLOTENAMTQ-ENARCFGSUVQL---QCSVVVGAPOEIVAANQSGSYQC--DYST 54
Db 30 YNVLPEAKIFSGPSSEQGYAVQOQINPKGNWLLVSGPWSGFPENRMGDYKCPVDLST 89
Qy 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGLC 107
Db 90 ATCEKLNQTSIPNVTETMKNLSGLILTRNMGTFGLTCGPLWAOCCGNYQYTTGVC 149
Qy 108 FLFGSNLRQPKPPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEWSTVMQK-- 165
Db 150 SDISPDF-QLSASFSPATQPCPSL-IDVVVWCDSENSIYPWD--AVKNFLEKFVQGLDIG 205
Qy 166 KSKTLFSLMQYSEERPHFTKEFQNNPNRSLIKPITQLLG-RTHATGRKVVURELFN 224
Db 206 PTKTQVGLIQVANNPRVVFNLTNYKTEEMIVATSSQSYGGDLTNTFGAIQYARKYAS 265
Qy 225 ITNGARKNAFKILFLTDGKFGDPLGVEDVIPELDREGVIRYIGV-----GDAFRSEK 279
Db 266 AASGRRSATKVMVVVTGDGESH-DGSMKAVIDQCNDHNILRFGIAVLGYLNRNALDTKN 324
Qy 280 SRQELNTVASKPPRPHVQINNFEALKTIONLREKIFAIEGTOTGSSSFEHEMSQEGF 339
Db 325 LIKEIKATASIPTERYFNVNSDEAALLEKAGTLGEQIFSIEGTVQG-GDNFQEMSQVGF 383
Qy 340 SAAITSNCP--LLSTVGSYDAGGVFLYTSKEKSTFINMT--RVDSMDN-DAYLGYAAAI 394
Db 384 SADYSSQNDILMLGAVGAFGMSGTIVQKTSHTGHILFPKQAFDQILQDRNHSYILGSVAA 443
```

Search completed: November 25, 2003, 14:38:49
Job time : 17.9635 secs

```
Qy 395 ILNRVQSVLVLGAPRYOHIGLIVAMFRONTGMWESNANV-----KQTQIGAYFASGLCSV 448
Db 444 ISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNITVIOAHRGDQGSYFSGVLCV 498
Qy 449 DVDNSNSTDLVLIGAPHYYEOTR--CGOVSVCPLPRQARWOCDAVLCEQOPWREG 506
Db 499 DVDKOTITDVLGAGPMYSMDLKKBEGRVYLFITKEGILGQHO---FLSGPEGIENTRFG 555
Qy 507 AALTVLGDVNGDKLTDAICAPGEEDNRGAVYLFHGTSGSIGSPSHSQRIAGS--KLSPR 564
Db 556 SAIAALSDINMDGFNDVIVGSPLENQNSGAVIYNGHQT-IRTKYSQKILGSDGAFRSH 614
Qy 565 LQYFGQSLGGQDLTWGDLVLTGACQHVLLRSQPVLRVKAIMEFNPREFVARNVFEBCN 624
Db 615 LQYFGRSLDGYGLNGDSITDVSIGAFGVQVQLWSQSIADVAIEASTPEKI--TLVNKN 672
Qy 625 DOVVKGKEAGEVAVCLHVQKSTRDLREGIOISVTVTYDLALD-----SGRPHSAVNEIK 680
Db 673 AQII-----LKLCF-----SAKPRPTKQNNQVAIVNITLDADGFSRSTSGLPKNN 721
Qy 681 NSTRROTQVLGLTQTC--ETLKLQLPNCIEDPVSPIVLRNLNLSVLTPLSAFGLNLRPVL 738
Db 722 ERCLQKNVNVQAQSCPEHIYIQEPS---DVVNSLDLVDISLENPGTS-----PALE 772
Qy 739 EDAQRLFTALFPPEKNCNDNICODDLSITF----SFMSLQCLVVGCPREFNVTVTRND 794
Db 773 AYSETAKVFSIPPHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQKRLTFSVTLKKN 832
Qy 795 GEDSYRTQVTFPPFLDLSYRKVSTLQNSQRSWRLACESASST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTVEVTCVAAASQKSQVACDVGY 880
Qy 854 PIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEFQLELPLVKYAVMV 913
Db 881 PALKRQOQVTFINFDENLQ-NLQNASLSFQALSSEQENKADNLVNLKIPLLYDAEIH 939
Qy 914 VTSHGVS TKYLNFTASENTS RVNMQHOYQ-----VSNLQORSLPISLVFLV----- 958
Db 940 LT-RSTNINFEISSDGNVPSIV-HSFEDVGPKFIFSLKVGSPVSMATVIIHIPOYTK 997
Qy 959 -----PVRLNQTWIDRPOVTF-SENLSSTCHTKERLPSH 992
Db 998 KNPLMYLTGVQTDKAGDISCNADINPLKIGQT-----SSVSFKSENFR---HTKE----- 1045
Qy 993 SDFLAELKAPVWNCIAVCQRIQCOIPFFGIOEENFATLKGNSLSPDWIYKTSNHLLIV 1052
Db 1046 -----LNCRTASCNVTCWLKDVHMKGEYFVNVVTRINWGTFASSITFTVQLT 1093
Qy 1053 STABI-LFNDVSFTLLPGQAGFVRSQETKVEPEVENP-----LP--LIVGSSVGGLL 1104
Db 1094 AAAEINTYPEIYVI-----EDNTVTIPLMIMKPKDEKAEVPTGVIIGSIAGILL 1143
Qy 1105 LALITAAALYKLGFFKROYKDM 1125
Db 1144 LLALVAILWKLGFKKRYEKH 1164
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:13:09 ; Search time 13.5618 Seconds
(without alignments)
3547.268 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFOENARGFGQ.....FKRYKDMMSGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/iaa/5A COMB.pcp.*
- 2: /cgn2_6/prodata/iaa/5B COMB.pcp.*
- 3: /cgn2_6/prodata/iaa/6A COMB.pcp.*
- 4: /cgn2_6/prodata/iaa/6B COMB.pcp.*
- 5: /cgn2_6/prodata/iaa/6C COMB.pcp.*
- 6: /cgn2_6/prodata/iaa/6D COMB.pcp.*
- 7: /cgn2_6/prodata/iaa/6E COMB.pcp.*
- 8: /cgn2_6/prodata/iaa/6F COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5852	99.5	1153	1	US-08-173-497-3
2	5852	99.5	1153	1	US-08-286-889-3
3	5852	99.5	1153	1	US-08-485-618-3
4	5852	99.5	1153	1	US-08-362-652-3
5	5852	99.5	1153	2	US-08-605-672-3
6	5852	99.5	1153	2	US-08-482-293A-3
7	5852	99.5	1153	2	US-08-943-363-3
8	5852	99.5	1153	3	US-09-193-043-3
9	5852	99.5	1153	4	US-09-688-307A-3
10	5821.5	98.9	1152	2	US-08-476-062A-43
11	5821.5	98.9	1152	5	PCT-US96-01314-43
12	5821.5	98.9	1152	6	5424399-2
13	3459	58.8	1163	2	US-08-476-062A-44
14	3459	58.8	1163	5	PCT-US96-01314-44
15	3436	58.4	1163	1	US-08-173-497-4
16	3436	58.4	1163	1	US-08-286-889-4
17	3436	58.4	1163	1	US-08-485-618-4
18	3436	58.4	1163	1	US-08-362-652-4
19	3436	58.4	1163	2	US-08-605-672-4
20	3436	58.4	1163	2	US-08-482-293A-4
21	3436	58.4	1163	2	US-08-943-363-4
22	3436	58.4	1163	3	US-09-193-043-4
23	3436	58.4	1163	4	US-09-688-307A-4
24	3401	57.8	1161	1	US-08-173-497-2
25	3401	57.8	1161	1	US-08-286-889-2
26	3401	57.8	1161	1	US-08-485-618-2
27	3401	57.8	1161	1	US-08-362-652-2

28	3401	57.8	1161	2	US-08-605-672-2	Sequence 2, Appli
29	3401	57.8	1161	2	US-08-482-293A-2	Sequence 2, Appli
30	3401	57.8	1161	2	US-08-943-363-2	Sequence 2, Appli
31	3401	57.8	1161	3	US-09-193-043-2	Sequence 2, Appli
32	3401	57.8	1161	4	US-09-688-307A-2	Sequence 2, Appli
33	3385.5	57.5	1161	1	US-08-485-618-99	Sequence 99, Appli
34	3385.5	57.5	1161	2	US-08-605-672-99	Sequence 99, Appli
35	3385.5	57.5	1161	2	US-08-482-293A-99	Sequence 99, Appli
36	3385.5	57.5	1161	2	US-08-943-363-99	Sequence 99, Appli
37	3385.5	57.5	1161	3	US-09-193-043-99	Sequence 99, Appli
38	3385.5	57.5	1161	4	US-09-688-307A-99	Sequence 99, Appli
39	3224.5	54.8	1161	3	US-09-193-043-55	Sequence 55, Appli
40	3224.5	54.8	1161	4	US-09-688-307A-55	Sequence 55, Appli
41	3217.5	54.7	1161	1	US-08-485-618-55	Sequence 55, Appli
42	3217.5	54.7	1161	1	US-08-362-652-55	Sequence 55, Appli
43	3217.5	54.7	1161	2	US-08-605-672-55	Sequence 55, Appli
44	3217.5	54.7	1161	2	US-08-482-293A-55	Sequence 55, Appli
45	3217.5	54.7	1161	2	US-08-943-363-55	Sequence 55, Appli

ALIGNMENTS

RESULT 1

US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-173-497-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCPTI 60
17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCPTI 76
Db

QY 61 RLQVPEAVNNMNSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFGLFSGNLRRQPOK 120
DB 77 RLQVPEAVNNMNSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFGLFSGNLRRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFINITNGARKNAFKILVLL 240
DB 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFINITNGARKNAFKILVLL 256
QY 241 TDGEKFGDPLGYEDVIPEDLREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPEDLREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 316
QY 301 NFEALKTIONQLREKIFAIEGTQGTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIFAIEGTQGTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGGQVSVCLP 480
DB 437 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGGQVSVCLP 496
QY 481 PRGQARWQCDVAVLGEQOPWGRFGAALTVLGDVNGDKLTDVAITGAPCEEDNRGAVLYF 540
DB 497 PRGQARWQCDVAVLGEQOPWGRFGAALTVLGDVNGDKLTDVAITGAPCEEDNRGAVLYF 556
QY 541 HCTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HCTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRKVAIMEFNPVARNVFCNDQVVKGEAGVRVCHLVOKSTRDLRREGQIQSVVT 660
DB 617 PVLRKVAIMEFNPVARNVFCNDQVVKGEAGVRVCHLVOKSTRDLRREGQIQSVVT 676
QY 661 YDLALDGRPHSRAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPISVLRNLF 720
DB 677 YDLALDGRPHSRAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPISVLRNLF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPFKNGCNDNICODDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPFKNGCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPFLDLRYKVTSLQNSORSWRMLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSVRTQVTFPFLDLRYKVTSLQNSORSWRMLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVNHQYQVSNLQSRSLPISLVFLVPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVNHQYQVSNLQSRSLPISLVFLVPV 976
QY 961 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRQCDIP 1020
DB 977 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRQCDIP 1036
QY 1021 FFGIQBEFNATLKGNLSPDWYIKTSHNHLIIIVSTAEILFENDSVFTLLPQOGAFVRSQTET 1080
DB 1037 FFGIQBEFNATLKGNLSPDWYIKTSHNHLIIIVSTAEILFENDSVFTLLPQOGAFVRSQTET 1096
QY 1081 KVEPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMWSEGGPPGABEQ 1137
DB 1097 KVEPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMWSEGGPPGABEQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3
Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCCEPI 60
DB 17 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCCEPI 76
QY 61 RLQVPEAVNNMNSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFGLFSGNLRRQPOK 120
DB 77 RLQVPEAVNNMNSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFGLFSGNLRRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTFLSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFINITNGARKNAFKILVLL 240
DB 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFINITNGARKNAFKILVLL 256
QY 241 TDGEKFGDPLGYEDVIPEDLREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPEDLREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVFQIN 316
QY 301 NFEALKTIONQLREKIFAIEGTQGTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIFAIEGTQGTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376


```
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
QY 481 PRGORARWQCDVLYGEOGPGRFGAALTVLGDVNGDKLTVDVAIGAPGEENRGAVYLP 540
DB 497 PRGORARWQCDVLYGEOGPGRFGAALTVLGDVNGDKLTVDVAIGAPGEENRGAVYLP 556
QY 541 HGTSGSGISPSHSQRIAGSKLSRLOQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSRLOQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
DB 617 PVLVRKAIIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRVAVENETKNSRROTQVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSRVAVENETKNSRROTQVGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNGLRPVLAEDAORLFTALPFPEKNCNDNIQDDLSITFSMSLCLVVG 780
DB 737 SLVGTPLSAFNGLRPVLAEDAORLFTALPFPEKNCNDNIQDDLSITFSMSLCLVVG 796
QY 781 GPREFNVTVVRNDEGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDEGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKANTSENNWPRNTKTEF 900
DB 857 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKANTSENNWPRNTKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQVQVSNLQORSLSPLSLVLPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQVQVSNLQORSLSPLSLVLPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 FFGIQEENATLKNLSFDWYIKTSHNHLILVSTABILFNDVSTLLPGGAFVRSQTET 1080
DB 1037 FFGIQEENATLKNLSFDWYIKTSHNHLILVSTABILFNDVSTLLPGGAFVRSQTET 1096
QY 1081 KVEPPEVPNPLPLIVGSSVGLLLALITAAALYKLGFKRQYKDMMSBEGGPPGABPQ 1137
DB 1097 KVEPPEVPNPLPLIVGSSVGLLLALITAAALYKLGFKRQYKDMMSBEGGPPGABPQ 1153
```

RESULT 3

```
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQGRSLYQCDYSTGSCPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQGRSLYQCDYSTGSCPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPQLLAGCPVTHQTCSENTYVKGCLFLFGSNLRQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPQLLAGCPVTHQTCSENTYVKGCLFLFGSNLRQPOK 136
QY 121 FPEARLGCPOEDSIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTLFSLMOYSEEF 180
DB 137 FPEARLGCPOEDSIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVIGVGDAPFRSEKSRQELNVTASKPRDRHVFOIN 300
DB 257 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVGDAPFRSEKSRQELNVTASKPRDRHVFOIN 316
QY 301 NFEALKTQNLQREKIFAIEGTQTCSSSFEHEMSQEGFSAAITSGNPILLSTVGSYDMAG 360
DB 317 NFEALKTQNLQREKIFAIEGTQTCSSSFEHEMSQEGFSAAITSGNPILLSTVGSYDMAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
QY 481 PRGORARWQCDVLYGEOGPGRFGAALTVLGDVNGDKLTVDVAIGAPGEENRGAVYLP 540
DB 497 PRGORARWQCDVLYGEOGPGRFGAALTVLGDVNGDKLTVDVAIGAPGEENRGAVYLP 556
QY 541 HGTSGSGISPSHSQRIAGSKLSRLOQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ 600
```

Db 557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLRVKATMEENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
Db 617 PVLRVKATMEENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 676
Qy 661 YDLALDGRPHSRVAVFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 720
Db 677 YDLALDGRPHSRVAVFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDCGDSYRTQVTFPPDLDSYRKVSTLQNRORSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDCGDSYRTQVTFPPDLDSYRKVSTLQNRORSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTEF 916
Qy 901 QLELPVKAVYVMVTSHGVSFKYLNFTASENTSRVMOHQYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKAVYVMVTSHGVSFKYLNFTASENTSRVMOHQYQVSNLQGRSLPISLVFLVPV 976
Qy 961 RLNOVTIWDROPVTSSENLSSTCHTKERLPSSHDFLAELRKAPVNCISIAVCORIQCIP 1020
Db 977 RLNOVTIWDROPVTSSENLSSTCHTKERLPSSHDFLAELRKAPVNCISIAVCORIQCIP 1036
Qy 1021 FFGIOEENATLKGNSLFDWYIKTSHNHLLIIVSTAEILFNDSVFTLLPGQCAFVRSQTET 1080
Db 1037 FFGIOEENATLKGNSLFDWYIKTSHNHLLIIVSTAEILFNDSVFTLLPGQCAFVRSQTET 1096
Qy 1081 KVEPEFVNPLPIVSSVGGLLALLALITAALYKLGFKRQYKQKMMSEGGPPGAEPO 1137
Db 1097 KVEPEFVNPLPIVSSVGGLLALLALITAALYKLGFKRQYKQKMMSEGGPPGAEPO 1153

RESULT 4

US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshhall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Beet Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARGFGOSVVOLOQSRVVVVGAPQBIIVANQSGSLYQCDYSTGSCPI 60
Db 17 FNLDTENAMTFQENARGFGOSVVOLOQSRVVVVGAPQBIIVANQSGSLYQCDYSTGSCPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCTSENTYVKGCLCFPGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCTSENTYVKGCLCFPGSNLRQOPQK 136
Qy 121 FPEALRCGPOEDSDIAFLIDCGSGIIIPHDPRRMKEWTVNMEQLKSKTLPFLMOYSEEF 180
Db 137 FPEALRCGPOEDSDIAFLIDCGSGIIIPHDPRRMKEFVTVNMEQLKSKTLPFLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGLRKVYRELFINTGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGLRKVYRELFINTGARKNAFKILVVI 256
Qy 241 TDGSKFGDPLGYEDVPEADREGVIRVIVGVGDFAFRSEKSRQELINTASKPRDHVFOIN 300
Db 257 TDGSKFGDPLGYEDVPEADREGVIRVIVGVGDFAFRSEKSRQELINTASKPRDHVFOIN 316
Qy 301 NFEALKTIQNLREKIPAEIGTQTGSSSSFEHMSQGFSAATISNGPILLSVTSYDWAG 360
Db 317 NFEALKTIQNLREKIPAEIGTQTGSSSSFEHMSQGFSAATISNGPILLSVTSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGHWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTREGGVSVCP 480
Db 437 QNTGHWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTREGGVSVCP 496
Qy 481 PRGORARWQCDALVYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDALVYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLRVKATMEENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
Db 617 PVLRVKATMEENPREVARNVFECNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 676
Qy 661 YDLALDGRPHSRVAVFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 720
Db 677 YDLALDGRPHSRVAVFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDCGDSYRTQVTFPPDLDSYRKVSTLQNRORSWRLACESASSTEV 840

Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPESEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPESEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 916
Qy 901 QLELPVKYAYVMVTSHGVSSTKYLNFTASNTSRVMQHQYQVSNLQORSLSPLSLVLPV 960
Db 917 QLELPVKYAYVMVTSHGVSSTKYLNFTASNTSRVMQHQYQVSNLQORSLSPLSLVLPV 976
Qy 961 RLNOTVIMDRPOVTFSENLSTCHTKERLPSSHDSFLAELRKAPVNCSTAVCORIQCDIP 1020
Db 977 RLNOTVIMDRPOVTFSENLSTCHTKERLPSSHDSFLAELRKAPVNCSTAVCORIQCDIP 1036
Qy 1021 FFGIOBEFNATLKNLSFDWYIKTSHNHLIVSTABEILFNDSVFTLLPGOGAFVRSQTE 1080
Db 1037 FFGIOBEFNATLKNLSFDWYIKTSHNHLIVSTABEILFNDSVFTLLPGOGAFVRSQTE 1096
Qy 1081 KVEPFEVNPPLIVGSSVGLLLALITAALYKLGFFKRYKQYKDMMSSEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGLLLALITAALYKLGFFKRYKQYKDMMSSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
: Sequence 3, Application US/08605672
: Patent No. 5817515
: GENERAL INFORMATION:
: APPLICANT: Gallatin, W. Michael
: APPLICANT: Van der Vieren, Monica
: TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
: NUMBER OF SEQUENCES: 103
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive, 6300 Sear Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/605,672
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/173,497
: FILING DATE: 23-DEC-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/286,889
: FILING DATE: 5-AUG-1994
: APPLICATION NUMBER: US 08/362,652
: FILING DATE: 21-DEC-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Williams Jr., Joseph A.
: REGISTRATION NUMBER: 38,659
: REFERENCE/DOCKET NUMBER: 27866/32684
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-474-6300
: TELEFAX: 312-474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1153 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein

US-08-605-672-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAQOEIVAAANQORSLYQCDYSTGSCDEP 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAQOEIVAAANQORSLYQCDYSTGSCDEP 76
Qy 61 RLQVPEAVNMSLGLSLAATTPPOLLAGCPVTHOTCSNTYVKGICFLFGLNLRQOPQK 120
Db 77 RLQVPEAVNMSLGLSLAATTPPOLLAGCPVTHOTCSNTYVKGICFLFGLNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKKSKTFLSLMQYSEBF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKKSKTFLSLMQYSEBF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAKILFL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAKILVVI 256
Qy 241 TDGEXFGDPLGYEDVI PELDREGVIRYVIGVGDAFRSEKSRQELANTVASKPRDHVFQIN 300
Db 257 TDGEXFGDPLGYEDVI PEADREGVIRYVIGVGDAFRSEKSRQELANTVASKPRDHVFQIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSEGEFSAITSNGLPSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSEGEFSAITSNGLPSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICGAPHYETOTRGQVSCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICGAPHYETOTRGQVSCPL 496
Qy 481 PRQQRARWQCDALYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
Db 497 PRQQRARWQCDALYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTVAQAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTVAQAQGHVLLRSQ 616
Qy 601 PVLRYKAIMFENPREVARNVFECNDQVYKGEAGEVRVCLHVOKSTRDLRREGIQSVVT 660
Db 617 PVLRYKAIMFENPREVARNVFECNDQVYKGEAGEVRVCLHVOKSTRDLRREGIQSVVT 676
Qy 661 YDLALDSDGRPHSRAVFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVSPVLRLNF 720
Db 677 YDLALDSDGRPHSRAVFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVSPVLRLNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLFTALPPPEKNCNDNICQDDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLFTALPPPEKNCNDNICQDDLSITFSFMSLCLVVG 796
Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPESEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPESEVFNITFDVDSKASLGNKLLKXANTVSENNMPTNKTEF 916
Qy 901 QLELPVKYAYVMVTSHGVSSTKYLNFTASNTSRVMQHQYQVSNLQORSLSPLSLVLPV 960
Db 917 QLELPVKYAYVMVTSHGVSSTKYLNFTASNTSRVMQHQYQVSNLQORSLSPLSLVLPV 976
Qy 961 RLNOTVIMDRPOVTFSENLSTCHTKERLPSSHDSFLAELRKAPVNCSTAVCORIQCDIP 1020
Db 977 RLNOTVIMDRPOVTFSENLSTCHTKERLPSSHDSFLAELRKAPVNCSTAVCORIQCDIP 1036

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.58; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.28; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFGQSVVQLOGSRVVVVGAPQEIIVAAORGSLYQCDYSTGCEPI	60
Db	17	FNLDTENAMTFQENARGFGQSVVQLOGSRVVVVGAPQEIIVAAORGSLYQCDYSTGCEPI	76
Qy	61	RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFGLGSLNLRQOPQK	120
Db	77	RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFGLGSLNLRQOPQK	136
Qy	121	FPALRCPCPEDSDIAFLIDGSGSIIPHDPRMKKEWSTVMEQIKKSKTLFSLMQYSEEF	180
Db	137	FPALRCPCPEDSDIAFLIDGSGSIIPHDPRMKKEWSTVMEQIKKSKTLFSLMQYSEEF	196
Qy	181	RHITFKFQNNPNRSLIKPITQLGRTHATGLRKVVRELFTNITGARKNAFKILFLL	240
Db	197	RHITFKFQNNPNRSLIKPITQLGRTHATGLRKVVRELFTNITGARKNAFKILFLL	256
Qy	241	TGKEKFDPLGYEDVPELDREGVIRYVIGVGDAFRSEKSRQELNVTASKPPRDHVFQIN	300
Db	257	TGKEKFDPLGYEDVPELDREGVIRYVIGVGDAFRSEKSRQELNVTASKPPRDHVFQIN	316
Qy	301	NFALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAATNSGPLLSTVGSVDWAG	360
Db	317	NFALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAATNSGPLLSTVGSVDWAG	376

Qy	361	GVFLYTSKEKSTFINNTRVDSNDKNDAYLGAAAAIILNRNVQSLVLCAPRYQHIGLVAMFR	420
Db	377	GVFLYTSKEKSTFINNTRVDSNDKNDAYLGAAAAIILNRNVQSLVLCAPRYQHIGLVAMFR	436
Qy	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGQGVSCPL	480
Db	437	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGQGVSCPL	496
Qy	481	PRGORARWOCDAVLYGEOGQPGWGRFGAALTVLVDGVDNGDKLTQVAICAPGEEDNRGAVYLF	540
Db	497	PRGORARWOCDAVLYGEOGQPGWGRFGAALTVLVDGVDNGDKLTQVAICAPGEEDNRGAVYLF	556
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRQ	600
Db	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRQ	616
Qy	601	PVLRVKAIWEPNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT	660
Db	617	PVLRVKAIWEPNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT	676
Qy	661	YDLALDSGRPHSRVAFENETKSTRROTQVLGLTQTCETIKLQLPNCIEDPVSPIVLRNLF	720
Db	677	YDLALDSGRPHSRVAFENETKSTRROTQVLGLTQTCETIKLQLPNCIEDPVSPIVLRNLF	736
Qy	721	SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFPEKNCNDNICDDLSITFSFMSLDCLVWG	780
Db	737	SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFPEKNCNDNICDDLSITFSFMSLDCLVWG	796
Qy	781	GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLONQORSQSWRLACESASTEV	840
Db	797	GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLONQORSQSWRLACESASTEV	856
Qy	841	SGALKSTSCSINHPIFPENSEVTENITFDVDSKASIGNKLLKANKYTSNNMPRTNKTET	900
Db	857	SGALKSTSCSINHPIFPENSEVTENITFDVDSKASIGNKLLKANKYTSNNMPRTNKTET	916
Qy	901	QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSLSLFLVPV	960
Db	917	QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSLSLFLVPV	976
Qy	961	RLNQTVINDRPPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1020
Db	977	RLNQTVINDRPPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1036
Qy	1021	FFGIOEENATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGOGAFVRSOTET	1080
Db	1037	FFGIOEENATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGOGAFVRSOTET	1096
Qy	1081	KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPO	1137
Db	1097	KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPO	1153

RESULT 8

US-09-193-043-3
Sequence 3, Application US/09193043
Patent No. 6251395
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6251395el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03

```
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match      99.5%; Score 5852; DB 3; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAANQSGSLVQCDSYSGSCEPI 60
Db 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAANQSGSLVQCDSYSGSCEPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFPLFSGNLROQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFPLFSGNLROQPOK 136

Qy 121 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQKKSTLFLSMQYSEEF 180
Db 137 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQKKSTLFLSMQYSEEF 196

Qy 181 RIHFTFKFQNNPRSLIKPITOLLGRTHATGLRKVVRELFINITNGARKNAFKILPLL 240
Db 197 RIHFTFKFQNNPRSLIKPITOLLGRTHATGLRKVVRELFINITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGVGDAFRSEKSRQELNLTIVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRVYVGVGDAFRSEKSRQELNLTIVASKPPRDHVFQIN 316

Qy 301 NFEALKTIONQLREKIFAIEGTQTCSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQLREKIFAIEGTQTCSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILNRVOSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILNRVOSLVLGAPRYQHIGLVAMFR 436

Qy 421 QNTGWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTGRGGQSVCP 480
Db 437 QNTGWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTGRGGQSVCP 496

Qy 481 PRGQARWOCDAVLGEOGPWGRFGAALTVLGDVNGDKLTDVAITGAPGEEDNRCAGVYLF 540
Db 497 PRGQARWOCDAVLGEOGPWGRFGAALTVLGDVNGDKLTDVAITGAPGEEDNRCAGVYLF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 616

Qy 601 PVLRVKAIMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

Qy 661 YDLALDSGRPHSAFNETKSTRQTVLGLTQTCETLKLQPNCEIDPSPVILRLNF 720
Db 677 YDLALDSGRPHSAFNETKSTRQTVLGLTQTCETLKLQPNCEIDPSPVILRLNF 736

Qy 721 SLVGTPLSAFGLNLRPVLAEDAORLFTALPPEKNGCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAORLFTALPPEKNGCNDNI CODDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPLDLISYKYSTLQNRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPLDLISYKYSTLQNRQSRWLACESASSTEV 856

Qy 841 SGALKSTCSINHPIFPENSEVTFNITPDVDSKASLGNKLLKANVTSNNMPTNKTEF 900
Db 857 SGALKSTCSINHPIFPENSEVTFNITPDVDSKASLGNKLLKANVTSNNMPTNKTEF 916

Qy 901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVNHQYQVSNLQSRSLPISLVFLVPV 960
```

```
Db 917 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVNHQYQVSNLQSRSLPISLVFLVPV 976

Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIOCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIOCDIP 1036

Qy 1021 FFGIQBEFNATLKNLSFDWYIKTSHNHLIIVTAEILFNDVSFTLLPGQGAFFVRSQTET 1080
Db 1037 FFGIQBEFNATLKNLSFDWYIKTSHNHLIIVTAEILFNDVSFTLLPGQGAFFVRSQTET 1096

Qy 1081 KVBPFFVNPPLPIVSGSVGGLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1137
Db 1097 KVBPFFVNPPLPIVSGSVGGLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1153

RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688.307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match      99.5%; Score 5852; DB 4; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAANQSGSLVQCDSYSGSCEPI 60
Db 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAANQSGSLVQCDSYSGSCEPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFPLFSGNLROQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFPLFSGNLROQPOK 136

Qy 121 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQKKSTLFLSMQYSEEF 180
Db 137 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQKKSTLFLSMQYSEEF 196

Qy 181 RIHFTFKFQNNPRSLIKPITOLLGRTHATGLRKVVRELFINITNGARKNAFKILPLL 240
Db 197 RIHFTFKFQNNPRSLIKPITOLLGRTHATGLRKVVRELFINITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGVGDAFRSEKSRQELNLTIVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRVYVGVGDAFRSEKSRQELNLTIVASKPPRDHVFQIN 316

Qy 301 NFEALKTIONQLREKIFAIEGTQTCSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
```

```
Db 317 NFALKTIONQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLLSTVGSYDMAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAIAAILRNVRQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAIAAILRNVRQSLVGLAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVCP 496
Qy 481 PRQORARWOCDAVLGYEQGPWGRFGAALTGLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQORARWOCDAVLGYEQGPWGRFGAALTGLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLVRKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKQSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAVNETKQSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFENLRPLVAEDAQLFTALPFEXKNCNDNICQDDLITFTFSMSLDCLVVG 780
Db 737 SLVGTPLSAFENLRPLVAEDAQLFTALPFEXKNCNDNICQDDLITFTFSMSLDCLVVG 796
Qy 781 GPREFNVTVTRNDGSDSYRTQVTFPPFDLSYRKVSTLQNRQSRQSWRLACSSASSTEV 840
Db 797 GPREFNVTVTRNDGSDSYRTQVTFPPFDLSYRKVSTLQNRQSRQSWRLACSSASSTEV 856
Qy 841 SGALKSTCSINHPISPENSEVFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTEF 900
Db 857 SGALKSTCSINHPISPENSEVFNITFDVDSKASLGNKLLKXANTYSENMPRTNKTEF 916
Qy 901 QLELPVKYAYVMVTSHGVSSTKYNFTASENTSRVMQHOYQVSNLQORSIPISLVELVPV 960
Db 917 QLELPVKYAYVMVTSHGVSSTKYNFTASENTSRVMQHOYQVSNLQORSIPISLVELVPV 976
Qy 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQC 1020
Db 977 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQC 1036
Qy 1021 FFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGOGAFVRSQTET 1080
Db 1037 FFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGOGAFVRSQTET 1096
Qy 1081 KVPEFVFNPLPLIVGSSVGLLLALITAAALKYKLGFFKROYKDMMSSEGGPPGAE 1137
Db 1097 KVPEFVFNPLPLIVGSSVGLLLALITAAALKYKLGFFKROYKDMMSSEGGPPGAE 1153

RESULT 10
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; Zip: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
/
/
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: Windows95
/ SOFTWARE: FastSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/476,062A
/ Filing DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/216,081
/ Filing DATE: 21-MAR-1994
/ APPLICATION NUMBER: 07/637,830
/ Filing DATE: 04-JAN-1991
/ APPLICATION NUMBER: 07/539,842
/ Filing DATE: 18-JUN-1990
/ APPLICATION NUMBER: 07/212,573
/ Filing DATE: 28-JUN-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Freeman, John W.
/ REGISTRATION NUMBER: 29,066
/ REFERENCES/DOCKET NUMBER: 00786/068003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 43:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1152 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: Internal
/ US-08-476-062A-43

Query Match 98.9%; Score 5821.5; DB 2; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPT 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPT 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCPVTHQTCSENTYVKGCLFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCPVTHQTCSENTYVKGCLFLFGSNLRQPOK 136
Qy 121 FPEARLGCPOEDSDIAFLIDGSGSIIPHDFRMKEMWSTVMEQLKKSKTLFSLMYSEEF 180
Db 137 FPEARLGCPOEDSDIAFLIDGSGSIIPHDFRMKEMWSTVMEQLKKSKTLFSLMYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGRLKRVVRELNFNITNGARKNAFKILV 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGRLKRVVRELNFNITNGARKNAFKILV 256
Qy 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVIGVGDAFRSEKSRQELNVTASKPRDHVFN 300
Db 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVIGVGDAFRSEKSRQELNVTASKPRDHVFN 316
Qy 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLLSTVGSYDMAG 360
Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHEMSQEGFSAAITSNGPLLLSTVGSYDMAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAIAAILRNVRQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAIAAILRNVRQSLVGLAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVCP 496
Qy 481 PRQORARWOCDAVLGYEQGPWGRFGAALTGLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQORARWOCDAVLGYEQGPWGRFGAALTGLDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
```



```
Db 556 HGTSGGSPSHSQRAGKSLSPRLQYFGQSLSCGQDLTMDGLVDLTVGAQGHVLLRSQ 615
Qy 601 PVLVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLEGGQIQSVVT 660
Db 616 PVLVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLEGGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 735
Qy 721 SLVGTPLSAFCNLRPLVLAEDAQRLLFTALFPPEKCGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFCNLRPLVLAEDAQRLLFTALFPPEKCGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 915
Qy 901 QLELPVKAVVMVTSHGVSNTKYNFTASENTSRVMQHOYVSNLQORSLSPLSVFLVPV 960
Db 916 QLELPVKAVVMVTSHGVSNTKYNFTASENTSRVMQHOYVSNLQORSLSPLSVFLVPV 975
Qy 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVCSIAVCQRIQCDIP 1020
Db 976 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVCSIAVCQRIQCDIP 1035
Qy 1021 PFGIOEBENATLKNLSFDWIKTSHNHLLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
Db 1036 PFGIOEBENATLKNLSFDWIKTSHNHLLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1095
Qy 1081 KVEPEVNPPLIIVGSSVGGLLIATLALITAAALYKLGPFKROYKDMWSSGGPPGAEPO 1137
Db 1096 KVEPEVNPPLIIVGSSVGGLLIATLALITAAALYKLGPFKROYKDMWSSGGPPGAEPO 1152
```

RESULT 11

```
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555x
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
```

```
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43
```

```
Query Match 98.9%; Score 5821.5; DB 5; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;
```

```
Qy 1 FNLDTENAMTFQSNARGFGQSVVQLQGSRRVWVGAPOEIVAAORGSLYQCDYSTGSGCEPI 60
Db 17 FNLDTENAMTFQSNARGFGQSVVQLQGSRRVWVGAPOEIVAAORGSLYQCDYSTGSGCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHQTCSENTYVKGCLFLPGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHQTCSENTYVKGCLFLPGSNLRQOPQK 136
Qy 121 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKWVSTVMEOLKXSKTLFSLMQVSEEF 180
Db 137 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKWVSTVMEOLKXSKTLFSLMQVSEEF 196
Qy 181 RIHFTFKFONNPNRSLIKPITQLLGRTHATGLRKVRELNIITNGARKNAFKILFL 240
Db 197 RIHFTFKFONNPNRSLIKPITQLLGRTHATGLRKVRELNIITNGARKNAFKILFL 256
Qy 241 TDGEKFGDPLGYEDVIPEDLREGVIRYVIGVGAFAFRSEKSRQELNVTASKPPRDHVPQIN 300
Db 257 TDGEKFGDPLGYEDVIPEDLREGVIRYVIGVGAFAFRSEKSRQELNVTASKPPRDHVPQIN 316
Qy 301 NFBAKTIQNLREKIPAIEGTQTGSSSSFEHEMSQGFSAAITNSGPELLSTVGSYDWAG 360
Db 317 NFBAKTIQNLREKIPAIEGTQTGSSSSFEHEMSQGFSAAITNSGPELLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 480
Db 437 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 496
Qy 481 PRGARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAAGEEDNRGAYLP 540
Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAAGEEDNRGAYLP 555
Qy 541 HGTSGGSPSHSQRAGKSLSPRLQYFGQSLSCGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 HGTSGGSPSHSQRAGKSLSPRLQYFGQSLSCGQDLTMDGLVDLTVGAQGHVLLRSQ 615
Qy 601 PVLVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLEGGQIQSVVT 660
Db 616 PVLVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLEGGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQLPNCIEDPSPVILRLNF 735
Qy 721 SLVGTPLSAFCNLRPLVLAEDAQRLLFTALFPPEKCGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFCNLRPLVLAEDAQRLLFTALFPPEKCGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 915
```


Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMOHQYQVSNLQSRSLPISLVFLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMOHQYQVSNLQSRSLPISLVFLVPV 975
Qy 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 976 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1035
Qy 1021 FFGIOEBFNATLKNLSDFWIKTSHNHLITVSTAEILFNDSVFTLLPGQAGAFVRSQTET 1080
Db 1036 FFGIOEBFNATLKNLSDFWIKTSHNHLITVSTAEILFNDSVFTLLPGQAGAFVRSQTET 1095
Qy 1081 KVEPFEPVNPPLIVGSSVGLLALITAAALYKLGFFKQYKQDMSEGGPPGAEPQ 1137
Db 1096 KVEPFEPVNPPLIVGSSVGLLALITAAALYKLGFFKQYKQDMSEGGPPGAEPQ 1152

RESULT 12
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO:2
; LENGTH: 1152
5424399-2

Query Match 98.9%; Score 5821.5; DB 6; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTQENARFGQSVOLQGSRVVVGAPQEIIVAAANQSGSYQCDSYSGCEPI 60
Db 17 FNLDTENAMTQENARFGQSVOLQGSRVVVGAPQEIIVAAANQSGSYQCDSYSGCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTQTCSENTYVKGCLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTQTCSENTYVKGCLCFLFGSNLRQOPQK 136
Qy 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTFLSLMQYSEEF 180
Db 137 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTFLSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPNSRLIKDITOLLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNSRLIKDITOLLGRTHATGLRKVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGKFGDPLGVEDVPELDREGVIRYVIGVDGAFRSEKSRQELNTVASKPRDRHVFOIN 300
Db 257 TDGKFGDPLGVEDVPELDREGVIRYVIGVDGAFRSEKSRQELNTVASKPRDRHVFOIN 316
Qy 301 NFEALKTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 436
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTEQTRGGVSVCPPL 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTEQTRGGVSVCPPL 496
Qy 481 PRGORARWQCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540

Db 497 PRG-RARWQCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSQRAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTVGAGQHVLRLRSQ 600
Db 556 HGTSGSGISPSHSQRAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTVGAGQHVLRLRSQ 615
Qy 601 PVLVRKAIMEFNPREVARNVFECDNVQVKGKEAGEVRVCLVHOKSTRDRREGIOQSVVT 660
Db 616 PVLVRKAIMEFNPREVARNVFECDNVQVKGKEAGEVRVCLVHOKSTRDRREGIOQSVVT 675
Qy 661 YDLALDSGRPHSRVAVNETKNSTRTOVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAVNETKNSTRTOVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 735
Qy 721 SLVGTPLSAFNGNRPVLAEDAQRLFTALFPFENKCGNDNICQDDLSITTFMSLDCLVVG 780
Db 736 SLVGTPLSAFNGNRPVLAEDAQRLFTALFPFENKCGNDNICQDDLSITTFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSRWLACASSTEV 840
Db 796 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSRWLACASSTEV 855
Qy 841 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 915
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMOHQYQVSNLQSRSLPISLVFLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVMOHQYQVSNLQSRSLPISLVFLVPV 975
Qy 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 976 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1035
Qy 1021 FFGIOEBFNATLKNLSDFWIKTSHNHLITVSTAEILFNDSVFTLLPGQAGAFVRSQTET 1080
Db 1036 FFGIOEBFNATLKNLSDFWIKTSHNHLITVSTAEILFNDSVFTLLPGQAGAFVRSQTET 1095
Qy 1081 KVEPFEPVNPPLIVGSSVGLLALITAAALYKLGFFKQYKQDMSEGGPPGAEPQ 1137
Db 1096 KVEPFEPVNPPLIVGSSVGLLALITAAALYKLGFFKQYKQDMSEGGPPGAEPQ 1152

RESULT 13
US-08-476-062A-44
; Sequence 44: Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842

FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-062A-44

Query Match 58.8%; Score 3459; DB 2; Length 1163;
Best Local Similarity 60.9%; Pred. No. 1.1e-289;
Matches 687; Conservative 142; Mismatches 294; Indels 6

[illegible]

```
QY 1 FNLDTENAMTFOENARGFQSVVLOGSRVVGAPQEIIVAAANORGSLYQCDYSTGSCPEI 60
Db 20 FNLDTBELTAFRVDSAGFSDVVQYANSMWVVGAPQKITAAANTGGLYQCGYSTGACEPI 79
QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCLFGLSGNLROQPOK 120
Db 80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHCEGGRNMYLTGLCLLGLPT--QUTQR 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTLFSLMQVSEEF 180
Db 138 LPVSRQECPRQEDIVFLIDGSGSISSRNFAFMNFVRAVISQFQPSQFSLMQVSNKF 197
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFINITNGARKNAFKILFUL 240
Db 198 QTHFTFEFRRTSNPLSLASVHQLQCFYTTATAIQNVVHRLPHASYGARRDAIKLIVI 257
QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVGDFAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 258 TDGKKGSDSDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKLNDIAKSPQEHIFKVE 317
QY 301 NFPEALKTIONQLEKIFAIEGTOTGSSSPEHEMSQEGFSAAITSNGLLSTVGSYDMAG 360
Db 318 DFDALDKIQLEKIFAIEGTETTSSSSPELEWAQEGFSAVPTDGPVLGAVGSFTWSG 377
QY 361 GVFLYTSKEKSTFINTRVDSMDNDAYLGAAAAIILNRVQSVLVLGAPRYQHIGLVAMPR 420
Db 378 GAFLYPPNMSPTFINNSQENVDNRDYLGYSTELALWKGVQSVLVLGAPRYQHTGKAVIET 437
QY 421 QNTGMHESNANVGTOIGAFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQVSVCP 480
Db 438 QVSRQRMKAETVGTQIGSYFGASLCSVDVDTGSTDVLVIGAPHYYEOTRGQVSVCP 497
QY 481 PRQARWOCDAVLGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 498 PRQWR-RWNCDAVLGEQGPWGRFGAALTVLGDVNGDKLTDVIGAPGEENRGAVYLF 556
QY 541 HGTSGSISPHSQRIAGSKSLPRLOYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPSPISPHSQRIAGSKSLRLOYFGQSLGGQDLTQDGLVDLAVGARGQVLLRTR 616
QY 601 PVLRVKAIEMFNPREVARNVFECDQWVKGEAGEVRVCLHVOKSTRDLRREGIOISVT 660
Db 617 PVLVWVGWQFIPAEIPRSFAFECREQVWSEQTLVQSNICLYIDKRSKNLLGSRDQSSVT 676
QY 661 YDLALDSGRPHSRVAVNETKNSRRTOTVGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 677 LDALDPGLSLPRAATQETKNSRLSRVRLGLKAHCENFNLLPSCVBESVPTITLRLNF 736
QY 721 SLVGTPLSAFNLPRVLAEDAQRLLTALFPFKXNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 TLVGKPELLAFNRLPMLAALAQRYFTASLPFKXNCADHICQDNLGISFSFPGCLKSLVG 796
QY 781 GPREFNVTVVRNDGDSVRTQVTFPPLDLSVRKYSTIQLQNSQRSWKLACESASSTEV 840
Db 797 SNLELNAEVMVMDGSDSRTTTFHPAGLSRYRYVAEGQKQQLRSLHLTCDSPAVG-- 854
QY 841 SGALKSTSCSINHPIPEENSEVTENITTFDVSXASIGNKLLKANVTSENNMPTKTEF 900
Db 855 SQTWSTSCRINHPIFGGAQITPLATFDVSPKAVLDGRLLLTANVSSENNPTRTSKTTF 914
QY 901 QLELPVKYAVYVMTVSHGSTYKYNFTAS-ENTSRVMQHQYQVSNLGRSLPISLVFLVP 959
Db 915 QLELPVKYAVYTVVSSHEOFTKYNFSESEKESHVAMHRYQVNNLGRDQLPVSINFVWP 974
QY 960 VRLNQTVWDRPOVTFSENLSSTCHTKERLPSHSDFLAEURKAPVNCVNCIACQRCQDI 1019
Db 975 VELNQBAVMDVEVSPQNPQLRCSSEKIAAPPASDFLAHQKPNVLDCSIAGCLRPRCDV 1034
QY 1020 PFGIOEBFNATLUGNLSFDWYIKTSHNHLIIVSTAEILFNDSVFETLLPQGAFAVRSOTE 1079
Db 1035 PFSVQEBELDFTLUGNLSFGWVRQIIOKKVSVVSVVAEITFDTSVYSQLPQGAFAVRAQIT 1094
QY 1080 TKVEPPEVPNPLPIVGSVGGLLLLALITALYKLGFFKROYKDMWSE 1128

Db 1095 TVLEKYKVNPTPLIVGSSIGLLELLALITAVLYKVGFPRQYKEMME 1143

RESULT 15
US-08-173-497-4
; Sequence 4, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Marehall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCES/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-173-497-4

Query Match 58.4%; Score 3436; DB 1; Length 1163;
Best Local Similarity 60.6%; Pred. No. 1.1e-287;
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFQSVVLOGSRVVGAPQEIIVAAANORGSLYQCDYSTGSCPEI 60
Db 20 FNLDTBELTAFRVDSAGFSDVVQYANSMWVVGAPQKITAAANTGGLYQCGYSTGACEPI 79
QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCLFGLSGNLROQPOK 120
Db 80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHCEGGRNMYLTGLCLLGLPT--QUTQR 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWSTVMEQLKSKTLFSLMQVSEEF 180
Db 138 LPVSRQECPRQEDIVFLIDGSGSISSRNFAFMNFVRAVISQFQPSQFSLMQVSNKF 197
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFINITNGARKNAFKILFUL 240
Db 198 QTHFTFEFRRTSNPLSLASVHQLQCFYTTATAIQNVVHRLPHASYGARRDAIKLIVI 257
QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVGDFAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 258 TDGKKGSDSDYKDVIPMDAAGIIRYAGVGLAFQNRNSWKLNDIAKSPQEHIFKVE 317
QY 301 NFPEALKTIONQLEKIFAIEGTOTGSSSPEHEMSQEGFSAAITSNGLLSTVGSYDMAG 360
Db 318 DFDALDKIQLEKIFAIEGTETTSSSSPELEWAQEGFSAVPTDGPVLGAVGSFTWSG 377
QY 361 GVFLYTSKEKSTFINTRVDSMDNDAYLGAAAAIILNRVQSVLVLGAPRYQHIGLVAMPR 420
Db 378 GAFLYPPNMSPTFINNSQENVDNRDYLGYSTELALWKGVQSVLVLGAPRYQHTGKAVIET 437
QY 421 QNTGMHESNANVGTOIGAFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQVSVCP 480
Db 438 QVSRQRMKAETVGTQIGSYFGASLCSVDVDTGSTDVLVIGAPHYYEOTRGQVSVCP 497
QY 481 PRQARWOCDAVLGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 498 PRQWR-RWNCDAVLGEQGPWGRFGAALTVLGDVNGDKLTDVIGAPGEENRGAVYLF 556
QY 541 HGTSGSISPHSQRIAGSKSLPRLOYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPSPISPHSQRIAGSKSLRLOYFGQSLGGQDLTQDGLVDLAVGARGQVLLRTR 616
QY 601 PVLRVKAIEMFNPREVARNVFECDQWVKGEAGEVRVCLHVOKSTRDLRREGIOISVT 660
Db 617 PVLVWVGWQFIPAEIPRSFAFECREQVWSEQTLVQSNICLYIDKRSKNLLGSRDQSSVT 676
QY 661 YDLALDSGRPHSRVAVNETKNSRRTOTVGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 677 LDALDPGLSLPRAATQETKNSRLSRVRLGLKAHCENFNLLPSCVBESVPTITLRLNF 736
QY 721 SLVGTPLSAFNLPRVLAEDAQRLLTALFPFKXNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 TLVGKPELLAFNRLPMLAALAQRYFTASLPFKXNCADHICQDNLGISFSFPGCLKSLVG 796
QY 781 GPREFNVTVVRNDGDSVRTQVTFPPLDLSVRKYSTIQLQNSQRSWKLACESASSTEV 840
Db 797 SNLELNAEVMVMDGSDSRTTTFHPAGLSRYRYVAEGQKQQLRSLHLTCDSPAVG-- 854
QY 841 SGALKSTSCSINHPIPEENSEVTENITTFDVSXASIGNKLLKANVTSENNMPTKTEF 900
Db 855 SQTWSTSCRINHPIFGGAQITPLATFDVSPKAVLDGRLLLTANVSSENNPTRTSKTTF 914
QY 901 QLELPVKYAVYVMTVSHGSTYKYNFTAS-ENTSRVMQHQYQVSNLGRSLPISLVFLVP 959
Db 915 QLELPVKYAVYTVVSSHEOFTKYNFSESEKESHVAMHRYQVNNLGRDQLPVSINFVWP 974
QY 960 VRLNQTVWDRPOVTFSENLSSTCHTKERLPSHSDFLAEURKAPVNCVNCIACQRCQDI 1019
Db 975 VELNQBAVMDVEVSPQNPQLRCSSEKIAAPPASDFLAHQKPNVLDCSIAGCLRPRCDV 1034
QY 1020 PFGIOEBFNATLUGNLSFDWYIKTSHNHLIIVSTAEILFNDSVFETLLPQGAFAVRSOTE 1079
Db 1035 PFSVQEBELDFTLUGNLSFGWVRQIIOKKVSVVSVVAEITFDTSVYSQLPQGAFAVRAQIT 1094
QY 1080 TKVEPPEVPNPLPIVGSVGGLLLLALITALYKLGFFKROYKDMWSE 1128
```

Db 318 DFDALDKIQNLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377
QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRNQSOSLVLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINNSQENVDMDRSYLGYSTELALWKGVQSVLVLGAPRYQHIGKAVIFI 437
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSTDVLIGAPHYYEQTRGGQVSVCP 497
QY 481 PRQBARWQCDVAVLGEQOPWGRREGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 498 PRGMF-RWMCDAVLYGEQHPWGRREGAALTVLGDVNGDKLTDVWIGAPGEBEENRGAVYLF 556
QY 541 HGTSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVDLTGVAQCHVLLLSRQ 600
Db 557 HGVLGPSISPSHSORIASQSLSRLOYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616
QY 601 PVLRYKATMEFNPRVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLWGVSNQFIPABIPRPAEPCREQVSEQTLVQSNICLYIDKRSKNLLGSRDIQSSVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSSTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 LDALAPGRLSPRATFOETKRSLSRVVLGKKAHCENFNLLPSCVEDSVIPIILRLNF 736
QY 721 SLVGTPLSAFGNLRPLVAEDAORLFTALPPPEKNCNDNICODDISITFSFMSLDCLVVG 780
Db 737 TLVGLPLAFLNLRPLMALAORYETASLPFEKNCADHICQDNLGISFSFPLKSLLVG 796
QY 781 GPREFNVTVVNDGEDSVYRQVTFPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 SNLELNAEVMVWMDGEDSVGTTITFSHPAGLSRYVAEQKQQLRSLHLTC--CSAPVG 854
QY 841 SCALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLILKANVTSENNMPRTNKTEF 900
Db 855 SQGTWSTSCRINHILFRGGAQITFLATFDVSPKAVGLDRKLLIANVSSENNIPRTSKTIF 914
QY 901 QLELPVKYAVYVMTVSHGYSTKYLNFAS-ENTSRVMOHQYQVSNLQORSPLISLVFLVP 959
Db 915 QLELPVKYAVYIVVSHEQFTKYLNFSEEEKESHVAMHYQVNNLQORDLPVSNFWVP 974
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDI 1019
Db 975 VELNQEAVMMDVEVSHQPQPSLRCSSEKIPASDPLAHIQKNPVLDCSIAGCLFRCDV 1034
QY 1020 PFFGIOEENATLKNLSFDWIKTSHNHLIIVSTAELFNDVSFTLLPGQCAFVRSOTE 1079
Db 1035 PSFSVQOEELDFTLKGNLSFGWVRQILQKKVSVVSAEIIFDTSVYSQLPQGEAFMRAOTI 1094
QY 1080 TKVEPPEFVNPLIVGSSVGLLLILALITAAALYKLGFFKROYKDMWSE 1128
Db 1095 TVLEKYKVNPIPLIVGSSIGLLLLALITAVLYKVGFFKROYKEMWEE 1143

Search completed: November 25, 2003, 14:23:17
Job time : 17.5618 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 14.9579 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481b-5

Perfect score: 5876

Sequence: 1 FNLDTEAMTFQENARGFGQ.....FKRQYKDMMSGGPPGAE PQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	5862	99.8	1153	1	RWHUB cell surface glyco
2	4470	76.1	1153	2	S00551 leukocyte surface
3	3479	59.2	1163	1	RWHUB cell surface glyco
4	1533.5	26.4	1170	2	S03308 cell surface glyco
5	1538.5	26.2	1163	2	F56126 lymphocyte fuction
6	1153	19.6	1179	2	A53213 integrin alpha-E c
7	1103.5	18.8	1151	2	A45226 integrin alpha-1 c
8	1085	18.5	1170	2	I45914 integrin alpha 2 s
9	1071	18.2	1178	2	S44142 VLA-2 protein homo
10	1068	18.2	1181	2	A33998 integrin alpha-2 c
11	1062	18.1	1180	2	A35854 integrin alpha-1 c
12	663	11.3	1039	2	A41131 lymphocyte-Peyer's
13	635	10.8	1038	2	S06046 integrin alpha-4 c
14	630	10.7	1035	2	I58409 integrin alpha-9 c
15	614.5	10.5	1041	2	T31437 integrin alpha cha
16	579.5	9.9	1054	2	JC7294 alphap integrin -
17	572.5	9.7	1051	2	A35761 cell surface glyco
18	567.5	9.7	1053	2	I55534 VLA-3 alpha subuni
19	555.5	9.5	1053	2	S44250 integrin alpha-5 c
20	542.5	9.2	1034	2	A36108 integrin alpha-V c
21	535	9.1	1044	2	T10050 integrin alpha-v c
22	532	9.1	1049	2	A27079 fibronectin recept
23	532	9.1	1073	2	B36429 integrin alpha-6 c
24	531.5	9.0	1072	2	A38457 integrin alpha-6 c
25	530.5	9.0	1051	2	A40021 integrin alpha-3 c
26	526	9.0	1048	2	A27421 integrin alpha-5 c
27	525.5	8.9	1091	2	A41543 integrin alpha-6 c
28	517	8.8	1044	2	T16516 integrin alpha-8 c
29	496.5	8.4	1146	2	S40311 integrin - fruit f

RESULT 1

RWHUB

cell surface glycoprotein CD11b precursor [validated] - human

N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C;Accession: A31108; A28915; A41600; A30892; A32218; A45526; A26091; 152567

R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, B.

A;Reference number: A31108; MUID:88315033; PMID:2457584

A;Accession: A31108

A;Molecule type: mRNA

A;Residues: 1-1153 <COR>

A;Cross-references: GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148

A;Note: part of this sequence was confirmed by protein sequencing

R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A;Reference number: A28915; MUID:88257215; PMID:2454931

A;Accession: A28915

A;Molecule type: mRNA

A;Residues: 1-499,501-965,'P',967-1153 <ARN>

A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA:

A;Note: the authors translated the codon TAC for residue 1129 as Thr

A;Note: part of this sequence, including the amino end of the mature protein, was con

R;Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally

A;Reference number: A41600; MUID:92073318; PMID:1683702

A;Accession: A41600

A;Molecule type: DNA

A;Residues: 1-9 <SHS>

A;Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215

R;Arnaout, M.A.; Renold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe

A;Reference number: A94193; MUID:88190151; PMID:2833753

A;Accession: A30892

A;Molecule type: mRNA

A;Residues: 917-1042 <AR2>

A;Cross-references: GB:M18044

R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recep

A;Reference number: A32218; MUID:89098893; PMID:2563162

A;Accession: A32218

A;Molecule type: mRNA

A;Residues: 9-1153 <HIC>

A;Cross-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

A>Note: part of this sequence was confirmed by protein sequencing
 R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 480-490, 1993
 A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
 n during evolution.
 A:Reference number: A46526; MUID:93123748; PMID:8419480
 A:Accession: A46526
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499,501-1153 <FLE>
 A:CROSS-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
 A>Note: the last three bases of intron 13, CAG, are included in some but not all mature
 A>Note: sequence extracted from NCBI backbone (NCBIP:121963)
 R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnsaut, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A:Reference number: A90664; MUID:87076671; PMID:3539202
 A:Accession: A26091
 A:Molecule type: protein
 A:Residues: 17-31 <PIB>
 A:Experimental source: granulocytes
 R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A>Title: Characterization of the myeloid-specific CD11b promoter.
 A:Reference number: I52567; MUID:92144986; PMID:1346576
 A:Accession: I52567
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-9 <RES>
 A:CROSS-references: GB:M84477; NID:g180184; PIDN:AAA51960.1; PID:g553219
 A:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
 C:Genetics:
 A:Gene: GDB:ITGAM; CR3A
 A:CROSS-references: GDB:120599; OMIM:120980
 A:Map position: 16p11.2-16p11.2
 A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
 F:17-1108/Domain: extracellular #status predicted <EXT>
 F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:465-473/Region: calcium/magnesium binding #status predicted
 F:530-538/Region: calcium/magnesium binding #status predicted
 F:593-601/Region: calcium/magnesium binding #status predicted
 F:1109-1134/Domain: transmembrane #status predicted <TM>
 F:1135-1153/Domain: intracellular #status predicted <INT>
 F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.8%; Score 5862; DB 1; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FNLDTENAMTFQENARGFQSVVQLQSGSRVVGAPQEIIVANQKSLYQCDYSTGSCPEI 60
 Db 17 FNLDTENAMTFQENARGFQSVVQLQSGSRVVGAPQEIIVANQKSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNLSGLSLAATTSPPQLLAGCPVTHQTCSENYVVKGLCFGLSGNLRLQKQK 120
 Db 77 RLQVPVEAVNLSGLSLAATTSPPQLLAGCPVTHQTCSENYVVKGLCFGLSGNLRLQKQK 136
 QY 121 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIIVMBQKSKTLFSLMOYSEEF 180
 Db 137 PPEALRGCPQEDSDIAFLVDGSGSIIPHDFRMKEFVSVMBEQKSKTLFSLMOYSEEF 196
 QY 181 RIHFTKEFQNNPNRSLTKPITQLLGRTHATGIRKVVRELFNITGARKNAFKILILI 240
 Db 197 RIHFTKEFQNNPNRSLTKPITQLLGRTHATGIRKVVRELFNITGARKNAFKILVVI 256
 QY 241 TDGEKFGDPLGVEDYIPEDAREGVIRYVIGVDGAFRSEKSKOELNTVASKPRDRHVQFN 300
 Db 257 TDGEKFGDPLGVEDYIPEDAREGVIRYVIGVDGAFRSEKSKOELNTVASKPRDRHVQFN 316

RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N:Alternate names: complement-3 receptor alpha chain

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999

C:Accession: S00551; I59078

R:Pytele, R.

EMBO J. 7, 1371-1378, 1988

A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with th

A:Reference number: S00551; MUID:88312584; PMID:3044779

A:Accession: S00551

A:Molecule type: DNA

A:Residues: 1-1153 <PVT>

A:CROSS-references: EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52983

A>Note: the authors translated the codon CAC for residue 569 as Gln

R:Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A:Reference number: 159078; MUID:86287312; PMID:2942940
A:Accession: 159078
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.1%; Score 4470; DB 2; Length 1153;
Best Local Similarity 73.9%; Pred. No. 2.5e-302;
Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;

QY 1 ENLDTENAMTFQENARGFGSVVQLOGSRVWVGAPOEIVAAORGSLYOCYSTGSCPEI 60
DB 17 ENLDTENAMTFQENARGFGSVVQLOGSRVWVGAPOEIVAAORGSLYOCYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVYKGLCFGLFGSNLRQOPK 120
DB 77 PLQVPPEAVNMSLGLSLAVSTVPQQLACGPTVHQNCKENTYVNGLCYLFSGNLRPPQ 136

QY 121 FPAALRGCEQSDIAFLVDGSGIIPHDPRRAKEFISTVMBOLKSKTLFSLMQVSEBF 180
DB 137 FPAALRGCEQSDIIVFLVDGSGIINNIDFQKMEFVSTVMEQFKSKTLFSLMQVSEBF 196

QY 181 RHFTFKFQNNPNPSLKPITOLLGRHTATGIRKVVRELPNITNGARKNAFKILILI 240
DB 197 RHFTFNDKRNPSHSPSHVPIKQLNGRTKTAGIRKVVRELPHKNGARENAAKTLVVI 256

QY 241 TDEKEFGDPLGYEDVPEADREGVIRVIGVGDAPRSEKSRQBLNTVASKPRDHFQIN 300
DB 257 TDEKEFGDPLGYEDVPEADREGVIRVIGVGDAPRSEKSRQBLNTVASKPRDHFQIN 316

QY 301 NFALNTIQNLREKIFALEGTGTGSSSFEHMSQEGFSAAITSNGPLSTVSGVDMAG 360
DB 317 NFALNTIQNLREKIFALEGTGTGSSSFEHMSQEGFSAAITSNGPLSTVSGVDMAG 376

QY 361 GVPLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNVSQSLVLCAPRYOHIGLVAMER 420
DB 377 GAPLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNVSQSLVLCAPRYOHIGLVAMER 436

QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQVSVCP 480
DB 437 ENFGTWEPTHTSKSGOIGSYFGASLCSVDMDADNTNLILIGAPHYYEOTRGQVSVCP 496

QY 481 PRGQARWQCDVLYGCEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRG-RARWQCEALLHGDQGHQHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

QY 541 HGTSGSGISPHSQRTAGSKLSPLOYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
DB 556 YGASIASLASHSASHRIIGAFHSPGLQYFGOSLGGQDLTMDGLMDLAVGAQGHVLLRSQ 615

QY 601 PVLVRVKAIMEFNPREVARNVFECNDQVWVGKEAGEVRVCLVHVKSTRDLRREGIOISVT 660
DB 616 PVLVRVKAIMEFNPREVARNVFECNDQVWVGKEAGEVRVCLVHVKSTRDLRREGIOISVT 675

QY 661 YDLALDSGRPHSAVFNETKNSTRROTQVLGTLTQCTETLKLQPLNCIEDPVSPIVLRLNF 720
DB 676 YDLALDPSVRSIRAFDETKNTRRTTQVFLGMLQKCTETLKLILPDCVDSDVSPILRLNY 735

QY 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALPFPEKNCNDNIQQDDLISITFSWISLDCLVG 780
DB 736 TLVGLPRLSGNLRPVLAEDAQRFLTALPFPEKNCNDNIQQDDLISITFSWISLDCLVG 795

QY 781 GPREFNVTVVRNDCGDSYRTQVTFPPFLDLSYRKVSTLQORSQSRML-ACESASSTE 839
DB 796 GQPDENMSTLRNDCGDSYRTQVTFPPFLDLSYRKVSTLQORSQSRML-ACESASSTE 855

QY 840 VSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANKVTSNNMPTNKTE 899
DB 856 GHGALKSTTWNINHPIFPANSEVTNITFDVDSKASLGNKLLKANKVTSNNMPTNKTE 915

QY 900 FOLELPVKYAVVMVVTSHGVSTKYNLFTASENTSRVMOHQYVSNLQORSLSPLSLVFLVP 959
DB 916 FOLELPVKYAVVMVVTSHGVSTKYNLFTASENTSRVMOHQYVSNLQORSLSPLSLVFLVP 975

QY 960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI 1019
DB 976 VQINNVTVDHPQVIFSQNLSSACHTEQKSPHNSFRDQLETPVLNCSVAVCKRIQCDL 1035

QY 1020 PFFGLOEENATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVTLPLPGQCAFVRSQTE 1079
DB 1036 PFFGLOEENATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVTLPLPGQCAFVRSQTE 1095

QY 1080 TKVEPFEVNPVPLIVSGVGLLLALITAAALYKLGPPFKROYKDMMSGGPPGAEPQ 1137
DB 1096 TKVEPFEVNPVPLIVSGVGLLLALITAAALYKLGPPFKROYKDMMSGGPPGAEPQ 1153

RESULT 3
RWHUIC

N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999
C:Accession: A35584; A35543; S00864
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J: Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A35584
A:Accession: A35584
A:Molecule type: DNA
A:Residues: 1-1163 <COR>
A:Note: this revision to the sequence from reference A35543 includes the carboxyl end
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J: Biol. Chem. 265, 2782-2788, 1990
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecul
A:Reference number: A35543; MUID:90153906; PMID:2303426
A:Accession: A35543
A:Molecule type: DNA
A:Residues: 1-834 <CO2>
A:Note: this sequence has been revised in reference A35584
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocy
A:Reference number: S00864; MUID:88166645; PMID:3327687
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755, 'L', 757-1163 <CO3>
A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119758; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat ho
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repea
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F:20-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61.89, 392.697, 735.899, 939.1050/Binding site: carbohydrate (Asn) (covalent) #status p

175 QYSEEPRIHFTFEFQNNPRSLIKPIITQLGRTHATGIRKVVRELFNITNGARKNAF 234
197 QFSTYKTEFDSDYKWKQDALLKHVKHMLLTNTFGAINVAVTEVFREELGARDPAT 256
235 KILILITDGKFGDPLYEDVIEADREGVIRVIGVDAPRSEKSRQELNVTASPPRD 294
257 KVLIIITDGB--ATDSNGIDAAKD-----IIRVIIGIKGHFKQESQETLHKFASKPASE 309
295 HVFOINFEALKTIONLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITNSGPLLSTVG 354
310 FVKILDTFEKLDLFTLQKLYVIEGTSKQDITSFNMLSSGGSADLSRGHNAVGVAG 369
355 SYDWAGVFP-LYTSKEKSTFINTRVDSMDNDAYLGAAA-IILNRVQSLVLGAPRYQH 412
370 AKDWAGGFLDKADLQDDTFLGNEPLTPEVRAGLYGTVTWLPSRQKTSLSAGAPRYQH 429
413 IGLVAMFR--QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYYEQT 470
430 MGRVLLFPQPGGGHNSQVQTIHGTQIGSYFGGELCGVDVDDGETELLIGAPLFYGEQ 489
471 RGGQVSVCLPRGQARAWQDAV--LYGEGQOPWGRFGAALTVLGDVNGDKLTDVAIGAP 528
490 RGGRVFIY-----QRRQLGHEEVSLEQDGPYPLGRFGEAITALTIDNGDLGVAVAGAP 544
529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDGLTMDGLVLT 588
545 LEE--QGAIVIFNGRHG-GLSPQSPQRIEQTQVLSGIQWFGRIHGVKDLGGDLADVA 601
589 GAQHVLLRSQVPLRVKVAIWEINPREVARNVFECNDQVY-KKKEAGEVRVCLVHVKSTR 647
602 GAESQMTVLSRRPVDVMTLMSFPAIPVHEVECSYSTSNMKKEGVNITICFQI-KSLY 660
648 DLREGOIQSVVYTDIALDSGRHSRAVENETKNSRTRQVQLGLTQTCETLKLQNPNCI 707
661 PQP-QGRVLNLTITLQDGHRTTRRGLFPGGRHELRNRNAV-TSMSCDTDFSFHPVPCV 718
708 EDVPSIVLRLNFSL---VGTPLS--AFGN-----LRPVLAEADAQLFTALFPFEKNCGN 757
719 QDLISPINVSLNFSLEEECTPRDQAGKDIPILRPSLSHSETWEI-----PFENKCGE 773
758 DNTCQDLSITFQMSLDCLVWGGRPFNVTVTRNDGEDSYRTQVTFPPFLDLSVRKVS 817
774 DKCEANLRFVSFSPARSRLRLTAFASLSVELSLNLEEDAYVWQDLHPHPPGLSPRKVE 833
818 TLQORSQSRWLACES--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKAS 875
834 ML---KPHSIOIPVSCBELPEERLLSRAL---SCNVSSPIFKAGHSVALQMFMTLVNSS 887
876 LGNKLILLKANVTSENN---MPRTNKTFOLELPVKYAVVMVVTSHGVSTKYLNFTASEN 931
888 WGSVELHANVTNNEDSDLEDNSATTI---IPILYPINILIQDQEDSTLYVSFTPKGP 944
932 TSRVMOHQYQV---SNLQORSLP-LSLVFLVPLRLNQTIVMDRPQVTFNSLSSTCHTK- 986
945 KIHQVXMYQVRIQPSIHDNIPTLEAVGVQPQPSSEGPITHQWSVQMEPPV--PCHYED 1002
987 -ERLPSHSD--FLAELRKAPVNCSTAVCORIQCDIPFFQIQEFENATLKNLSFPMYIK 1043
1003 LERLPDAEPCPLGALFRCPVW-----FROEILVQVITGTLVELGGEIE 1044
1044 TSHNHLIVSTAELFNDSVFTLLPGOGAFVRQOTETKVPPEFPVNPPLPLIVGSSVGLL 1103
1045 AS-SMFSLSLSGISFNSSKHFLHYGNSASL-AQVVMKVDDVYKQMLYLYVLSSGIGLL 1102
1104 LLALITAALYKLGFFRKQYKDDMMSEG-GPPGAEF 1136
1103 LLLLIFIVLYKVGFRENTLKEKMEAGRGVNGIP 1136

RESULT 5
I56126
lymphocyte fuction-associated molecule-1-alpha - mouse

C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Aug-1999
C:Accession: I56126
R:Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A:Reference number: I56126; MUID:91268576; PMID:2051027
A:Accession: I56126
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1163 <RES>
A:Cross-references: GB:M60778; NID:9198785; PIDN:AAA39426.1; PID:9198786
C:Genetics:
A:Gene: LFA-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hc
F;151-315/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 26.2%; Score 1538.5; DB 2; Length 1163;
Best Local Similarity 34.2%; Pred. No. 2.4e-98;
Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;
QY 1 ENLDTENAMTFQENA-RGFGQSVVOLQGSRRVVVGAPOEIVAAQNRGSLYOCDYSTGSCPE 59
DB 24 YNLDRTPQTSLAQGRHFGYQLIEDG-VVVGAPGE---GNTGGLYHCRTSSEFCQP 79
QY 60 IRLQVPEVAVNMNLSGLSLAATSPPOLACGPTVHOTCSENTYVVKGLCFGLFNSLRQQPQ 119
DB 80 VSLH-GSNHTSKVLTGTLATDAKGSLLACDPCLSRTCDQNTYLSGLCYLFPQSLGPM 138
QY 120 KPPEALRGCPQDSDIAFLVDGSGSII PHDPRAKEFISTVMEQLKKSKTFLSLMYSEE 179
DB 139 QNRPAQECMKGKVDLVFLFDGQSQSLDRKDFEKLFEFMKDMVRKLSNTSYQFAAVGFST 198
QY 180 FRIHFTFEKF-QNNNPRSLIKPIITOLLGRTHATGIRKVVRELFNITNGARKNAKIL 238
DB 199 CRTEFTFLDYVKNKNPVDLLGSPQMFLLTNTFRAINVYVAHVFKESGARDATKVLV 258

QY 239 LITDEKFGDPLGYEDVIEPADREG-----VIRYVIGVGAFAFRSEKSRQELNVTASKP 291
DB 259 IITDG-----EASDKGNISAAHITVYIIGIKHFSVQKQTLHIIFASEP 304
QY 292 PRDHVQINNFEALKTIONLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITNSGPLUS 351
DB 305 VEEFVKILDTFEKLDLFTLQRRYIAIEGTNRQDLTSFNMLSSSGISADLSKGHAVVG 364
QY 352 TVGSDVMAGVFP-LYTSKEKSTFINTRVDSMDNDAYLGAAA-IILNRVQSLVLGAPR 409
DB 365 AVGAKDMAGGFLDLREDLQAGFVQGEPLTSDVRGGYLGTVVAMWTSRSRPLAAGAPR 424
QY 410 YQHIGLVAMFR--QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYY 467
DB 425 YQHVQVLLFQAPAEAGRWNTQKTEGTQIGSYFGGELCSVDLDQDGEAELLIGAPLFF 484
QY 468 EOTRGQVSVCLPRGQARAWQDAVLYGEGQOPWGRFGAALTVLGDVNGDKLTDVAIGA 527
DB 485 GEORGGRVFTY---QRRQSLFEMVSELOQDGPYPLGRFGAAITALTIDNGDRLTDVAVGA 541
QY 528 PGEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDGLTMDGLVLT 587
DB 542 PLEE--QGAIVIFNGKPG-GLSPQSPQRIEQTQVLSGIQWFGRIHGVKDLGGDLADVV 598
QY 588 VQAQGHVLLRSQVPLRVKVAIMEFNPREVARNVFECNDQVVKGKEAG-EVRVCLHVQKST 646
DB 599 VQAGRVVLSRRPVDVMTLMSFPAIPVHEVECSYSTSNMKKEGVNITICFQI-KACFR 658
QY 647 RDLRGVQSVVYTDIALDSGRPHSRVAFNETKNSRTRQVQLGLTQTCETLKLQNPNC 706
DB 659 PQ--FOGRLANLSTYTLQDGHMRSGFLPDGSHLSGNTSITP-DKSCLDPHFHPFIC 715
QY 707 IEDPVSIVLRLNFSLV---GTPLSAFGN-LRPVLAEADAQRFTALFPPEKNCNDNICQ 762
DB 716 IQDLISPINVSLNFSLEEEGTPRDQGRAMQPILRPSIHTV-TKEIPPEKNCGEDKKCE 774

A;Accession: A45226
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1151

A;Experimental source: hepatoblastoma cell line HepG2
A;Note: Sequence extracted from NCBI backbone (NCBI:P124326)
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F;142-317/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 18.8%; Score 1103.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 4.3e-486; Indels 195; Gaps 44;
Matches 344; Conservative 212; Mismatches 486;

QY 1 FNLDTENAMTFQENARG-FQSGVYVQL---QGSVVVVGAPQEIIVAAANQORGLYOCDSYSGS 56
DB 1 FNVVKNMSTFSQVDMGTYVQYENEEGKWLIGSLVGPQPKRTGDVYKCPVGRGE 60
QY 57 CEP-IRLVQVVEA-----VMSLGLSLAATTPPQLLACGPTVHOTCSNTYVYKGL 106
DB 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGGFLACGPLYAYRCGLHLHTTGI 119
QY 107 CFLFGSNLRQOPKFEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLK- 165
DB 120 CSDVSPTFQVNVSIAP--VQECSTQ-LDIVIVLDGSGSIYPWDSVTA--FLNDLLKRMDI 174
QY 166 -KSKTLFLSMOYSEEFRIHFTFKFQNNPNRSLIKPITQLLGR-THTATGIRKVVREL 223
DB 175 GPKQTQGVIVQYGENVTHEFNLNKYSSTEVLVAANKIVORGRTWTALGTDTRAKEAF 234
QY 224 NITNGAKNAFKILILITDGEKDPGLGYEDVPEADREGVIRYVIGVGDAFR-----SE 278
DB 235 TEARGARRGVKVMVITDGEH-DNHLKKVQDCEDENIQFSAIILGYSYNRGNLSTE 293
QY 279 KSRQELNTVASKPRDHVFOINNFEALKTIONLRKIPAEIGTQTGSSSPHEMSQEG 338
DB 294 KFVEETKSIASEPTKHFNFVSDALVTIKTLGERIFALEATADQSAASFEMEMSQTG 353
QY 339 FSAAITSNGLLSTVSGYDMAGGVFLYTSKE-----KSTF-INWTRVDSMDNDAVLGYAA 392
DB 354 FSAHYQDWMLGAVGAYDNWGVVMQKASQIIIPRNTTFNVSTKNEPL-ASYLOYTV 412
QY 393 AILNRVRVSL-VLGAAPRQVHIGLVAMFRONTQGMESNANVKGTQIGAFGSLCSVDVD 451
DB 413 NSATASSGDVLYTAGOPRYNHTQVYIYRMEDCNILQTLSEQIGSYFGSILTTDID 472
QY 452 SNGSTDLVLIGAPHY-----YEOTR-GGVSVCLPRGQRARQCDAVLY 495
DB 473 KDSNTDILLVGAAPYMGTEKEEQKVYVVALNQTREYQMSLBPDKQCCSSRQHSCTT 532
QY 496 GEOGQPMW-RFGAALTVLGDVNGDKLTDVAIGAPGEDNRRGAVYLFHGTSGSGISPSHSQ 554
DB 533 ENKNEPCGARFGTAIAAVDNLNDGNDIVIGAPLDDHGGAVYIYHG-SGKTIKEYAQ 591
QY 555 RIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTVGAQGHVLLRSOPVLVRVKAIMEFNPR 614
DB 592 RIFSGDGGKTLKFPQSGIHEMDLNGDGLTDVIGLGGALFWSRDAVAVKVTMFEPN 651
QY 615 EVARNPECDNQVVKKEAG--EVRVCLHVQ-KSTRRLREGQIQSVVYDLALDSGRPH 671
DB 652 KNIQKKNCH--MEGKETVCINATVCFEVLKSKEDTIVEADLQ---YRVTLDSLRQI 704
QY 672 SRAVFNET-----KSTRQTOVLGTQTCETLKLQPLNCIEDPSPVILRLNFSLVGT 725
DB 705 SRSPFSGTQBRKVORNTITRKSEC-----TKHSFYMLDKRHDQDSVR---ITLDFNLT-D 755
QY 726 PLSAFGLNLRVLAEDAQRFTALFPPEKNGCNDNICODLSITFSFMSLDCLVGGPRE- 784
DB 756 PENG-----PVLDLSLNSVHEIYIPAKDCGNKEKICISLDSLHVATEKDLLIVRQNDK 810
QY 785 FNVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWRACESASSTEVSGAL 844
DB 811 FNVSLTVKNTKDSAYNTRTIVHVSPLNVFSGIEAIQKD-----SCSEN----- 853

845 KSTSCSINHPIPPENSEVTFNITFDVDSKASIGN-KLLLLKANVTSENMPRTNKTEFOLE 903
DB 854 HNTCKVGYPFLLRRGEMVTKILFQNTSYLNMENVTIYLSATSDEEPPETLSDNVNVIS 913
QY 904 LPVKYVYVYVTSHGVSFKYLNFTASENTSRVMQHQYQVSN-----LQORS-----L 950
DB 914 IPVKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNIFLYLIRKSGSPMPBEL 972
QY 951 PISLVF-----LVPVRLNQTVIWDPRQVTFSENLSSTCHTKE-----RLPS 991
DB 973 KLSISFPNMTSGYVPVLYPTGLSS-----SENANCRPHIFEDPFSINSKGKMT 1021
QY 992 HSDFLAEIRKAPVNCVSIACVQRIQCDIPFFQIQE-----EFNATLK 1033
DB 1022 STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILWKPTFIKSYFESSLNLTR 1078
QY 1034 GNLSPDWYKTHSHNHLILVSTAEIILFNDSVFTLLFGQAGFVRSQRTETKVEPEVPNPLPL 1093
DB 1079 GEL-----RSENASLVSSN-----OKRELAIQISKDGLFGRVPL 1114
QY 1094 --IVGSSVGLLLALITAAALYKLGFFKRYQKDMSE 1128
DB 1115 WVILLSAFAGLLMLLILALWKIGFFRPLKXKMEK 1151

RESULT 8
I45914
integrin alpha 2 subunit - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999
C;Accession: I45914
R;Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A;Title: Identification of putative ligand binding sites within the I-domain of integrin alpha 2 subunit of bovine platelet glycoprotein IIb/IIIa
A;Reference number: A54402; PMID:94193847; PMID:7511592
A;Accession: I45914
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1170 <KAM>
A;Cross-references: GB:I25886; NID:9439695; PIDN:AAB59255.1; PID:9439696
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F;161-336/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.5%; Score 1085; DB 2; Length 1170;
Best Local Similarity 27.6%; Pred. No. 8.6e-67;
Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFQ-ENARGFQSGVYVQL---QGSVVVVGAPQEIIVAAANQORGLYOC--DYST 54
DB 19 YNVGLPKAKIFSGPSSEQFGYAVQOPIPKGNWLLVGSWSPFKRMGDVYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTPPQLLACGPTVHOTCSNTYVYKGLC 107
DB 79 TTCEKLNLTQTSMSNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLWAOQSGSYVTTGVC 138
QY 108 FLFGSNLRQOPKQKFEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEOLK-- 165
DB 139 SDVSPDF-QLRSTFPAPVQTCF-SFIDVVVVCDESNIYPWD--AVNLFLEKPVQGLDIG 194
QY 166 KSKTLFSLMQYSEEFRIHFTFKFQNNPNRSLIKPITQLL-----GRTHATGIRKVVRE 221
DB 195 PIKTQMGILQYANPRVFNLTAFKSKD---EMIKATSTQTFQYGGDLTNTFRAIOYARDT 251
QY 222 LFNITNGARKNAFKILILITDGEKDPGLGYEDVPEADREGVIRYVIGV-----GDAPR 276
DB 252 AYSTRAGGRPGATKVMVVTGDGESH-DGSKLKAVIDQCNKDNILRFGIYAVLGLYLRNALD 310
QY 277 SKSRQELNTVASKPRDHVFOINNFEALKTIONLRKIPAEIGTQTGSSSPHEMSQ 336
DB 311 TKNLKEIKAIASIPTEHFHFNVSDEADLEKAGTIGEQIFSTIEGVQCG-GDNFQNMESQ 369
QY 337 EGFSAAIT--SNGPLLSTVSGYDMAGGVFLYTSKEKSTFINMT--RVSDMDN-DAYLGYA 391

Db 370 VGFAEYSPONNMLGAVGAYDWSGTGVVQKTPHGHILFSKQAEQILQDRNHSSYLGS 429
Qy 392 AAILLRNRVQSILVIGAPRYQHIGLVAMFRONTGMWESNANV-----KGTOIGAYFGASL 445
Db 430 VASISTGNSVHFVAGAPRANTYGGQVLYSVN-----ENGNVTVQSQGGDQIGYFGSVL 484
Qy 446 CSVDVDSNGSDLVILIGAPHYYEOTR--GGQVSVCPPLRGORARWQCDVAVLYGQGPWG 503
Db 485 CAVDVNKDTITDVLVAGAPMYNDLKEEGRVLYFTTKG-ILNWH--QFLEGPNGLENA 541
Qy 504 RFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVLFGHTSGSGISPSHSORAGS--KL 561
Db 542 RFGSAIALSDINMDGDFNDVIGSPLENQNSGAVIYNGHEGM-IRLAYSQKILGSDRAF 600
Qy 562 SPRLQYFGOSLGGDGLMDGLVDLTGCAQGHVLLRQOPVLRKAIEMFNPREVARNVF 621
Db 601 SSHLYQFGRSLDGYDLNGDSITDVSGAFQGVQVLWQSQSDADVSVDASFTPKXI--TLL 658
Qy 622 ECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDLALD-----SGRPHSRAVFN 677
Db 659 NKNAEI-----KLKLCF---SAKPRPTNQNNQVAIVYNTITDQFSSRVISRGLFK 707
Qy 678 ETKNSTRTQVGLGTQCE--TLKQLQPNCTEDPVSIVLRNPSL--VGTPLSAFQNL 733
Db 708 ENNERCLOKMTVISOARCSEYIIHQEPS---DIISPLNLCMNISLENPGT-----756
Qy 734 RPLVAEDAQRFTALFPPEKXGNDNICODDLSITF---SFMSLDCLVWGPREFNVTV 789
Db 757 NPALAYSETVKVFSIPHKCCGDDGVCLISDLVNVQQLPATQQOPPIVSNQNKLTFSV 816
Qy 790 TVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESAST-EVSGALKSTS 848
Db 817 QLKNNKESAYNTEIVVDSENLFF-----ASWSPVVDGTEVTCQIASSQKSVT 864
Qy 849 CSINHIPIPENSEVTFTNFTVDVDSKASLGNKLLKANVTSENMPRTNKTETFEQLPKY 908
Db 865 CNVGYPALKSQOQVTFITNFDNLQ-NLQNOASISFRALSESQEEENMADNSVNLKSLLY 923
Qy 909 AVYVMTVSHGVSTKYLNTFASNTSRVMQHOVQVSNLQOR-----SLPTSLVELV 958
Db 924 DAEIHIT-RSTNINYEVSLDGNSV-VHSE--DIGPKFISKVTGTVGVPVMA---976
Qy 959 PVRLNQTVIMDRPQVTFSEN---LSSTCHTKE-----RLPSHSDFLAE- 998
Db 977 -----SVIIHIPQVTKDKNPLMYLTGVHTDQAGDISCEAEINPLKIGOTSSTSSVSKSEN 1030
Qy 999 LRKAPVNCIAVCORIQCIDIFFGIGIEFNATLKNLSFDWIYKTSNHLIVSTAEI- 1057
Db 1031 PRHIKELNCRATSCSNIMCWLDRDQVGEYFLNVSTRIWNGTFAASTFTQVLTAAAEID 1090
Qy 1058 LFNDSVFTL-----LPGQGAFFRSQVETETKVEPFE-VPNPLPLIVGSSVGGLLLALITA 1110
Db 1091 TYNPQIVYEENTVTP-----LTMKPEKVEVPTGVIVGSVIAGILLALLVA 1140
Qy 1111 ALYKLGFFKRYOKDM 1125
Db 1141 ILWKLGFPPKRYEKM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Danjanovich, L.; Latze

submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagens and laminin adhesion but not

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A:Cross-references: EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473099
C:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
F;169-344/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.2%; Score 1071; DB 2; Length 1178;
Best Local Similarity 27.9%; Pred. No. 8.2e-66;
Matches 342; Conservative 208; Mismatches 487; Indels 188; Gaps 44;
Qy 1 FNLDTENAMTFQ-ENARGFGQSVVQL---CGSRVWVAGAPOEIVAAANQGRSLYQC--DYST 54
Db 27 YNVGLPGAKIFSGPSSEQFGYSVQQLTNPQGNWLLVGSPPMSGFENRMGDVYKCPVDLPT 86
Qy 55 GSECEPIRLQ-----VPVEAVNMSLGLSLAATSPOLLACQPTVHQTCESTYVYKGLC 107
Db 87 ATCEKLNQNSASISNVEIKTNMSLGLTLTRNPGTGGFLTCPLMAHQCGNYYATGIC 146
Qy 108 FLFGSNLRQPOQ---KFEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTVNEQL 164
Db 147 -----SDVSPDFQFLTSPAPVQACPSL-VDVVVVCDSESNLYP--WEAVNKLKVFVTGL 199
Qy 165 K--KSKTLFSLMQVSEFRIHFTKFEQNNPNRSLIKPITQLLG-RTHTATGIRKYVRE 221
Db 200 DIGPKKTQVALIQYANEPRIIFNLNDFETKEDVMVQATSETROHGGDLTNTFRAIEPARDY 259
Qy 222 LFNITNGARKNAFKILITDGEKFGDPLGYEDVPEADREGVIRYVIGV-----GDAFR 276
Db 260 AYSQTSQGRPGATKVMVVVTDGESH-DGSKLTKTVIOQCNDDEILRFGIAVLGYLNRNALD 318
Qy 277 SEKSRQELNVTASKPPRDHVFQINNFEALQTIQNLREKIPATEGTGTSSTSSPEHEMSQ 336
Db 319 TKNLKEIKAIKASTPTERYFFNVADAEALLEKAGTLGEQIFSEGTVOG-GDNFQMEWAQ 377
Qy 337 EGESA--AITSNGPLLSVGSYDMAGVFLYTSKEKSTFINMT--RYDSMN-DAYLGYA 391
Db 378 VGFSADYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIFPKQAFDQVLQDRNHSPLGYS 437
Qy 392 AAILNRVOSLVGAPRYQHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASLCS 447
Db 438 VAAISTEDGVHFGVAGAPRANTYGGQVLYSVN---QGNVTVIOSHRGDQIGSYFGSVLCS 494
Qy 448 VDVSNGSTDLVLIGAPHYYEOTR--GGQVSVCPPLRGORARWQCDVAVLYGQGPWGPRF 505
Db 495 VDVQKOTITDVLVAGATPTMNDLKEEGKVFYFTITAGILNQHO---FLGPEGTGNARF 551
Qy 506 GAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVLFGHTSGSGISPSHSORAGSKLSPR- 564
Db 552 GSAIATLSDINMDGDFNDVIGSPVENENSGAVIYNGHQGT-IRTKYSQKILGSDNGAFRR 610
Qy 565 -LQYFGOSLGGQDLTMDGLVDLTGCAQGHVLLRQOPVLRKAIEMFNPREVARNVPEC 623
Db 611 HLQPFGRSLDGYDLNGDSITDVSGALQVQLWSQSIADVAIEALFTP-----660
Qy 624 NDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDLALD---SGRPHSRAVFNET 679
Db 661 -DKITLLNKOKAITLKLCFRAEFPAGQNNQV--AILFMTLDADGHSSRVTSRGVFRN 717
Qy 680 KNSTRRTQVGLGTQTCET--LKLQPLNCTIEDPVSIVLRNLSFVLGTPLSAFGNLRPVL 737
Db 718 SERFLQNMVNVVEQKCSHHISIQKPS---DVVNPLDLRVDISLENPGTS-----PAL 768
Qy 738 AEDAQRFTALFPPEKXGNDNICODDLSI-----TFSFMSLCLVVGGRPFNFTV 788
Db 769 EAYSETVKVFSIPFKEGSDGICISDLILDVQQLPATQTSF-----IVSNQNKRLTFS 823
Qy 789 VTVRNDGEDSVYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESAST-EVSGALKAST 847
Db 824 VILKNRGESAYNTVVLAEFSLENLFF-----ASFMPVDGTEVTECEVSSSQKSV 871
Qy 848 SCSENHPIFPENSEVTFTNFTVDVDSKASLGNKLLKANVTSENMPRTNKTET--FQLELP 905
Db 872 TCDVGYPALKSEQQVTFITNFDNLQ-NLQNOAAINFOAFSESO--ETNKADNSVSLTIP 928
Qy 906 VKYAVYVMTVSHGVSTKYLNTFASNTSRVMQHOVQVSNLQOR-----SLPTSLV 955

```

Db 929 LLVDAELHLT-RSTNINFISSDENAPSVIK---SVEDIGPKPIFSLKVTAGSAPVMA 984
Qy 956 FLV-----PVRINQTVIMDRPQVTF-SENLS 980
Db 985 LVTHIQYTKENPLLYLTGIQTDQAGDISCTAEINPLKLPHTA-----PSVDFQENFR 1040
Qy 981 STCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQCDIPFFGIQEBFNATLKNLSFDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMAEYFINVTRVMNRT 1080
Qy 1041 YIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEKVEPPEFVNPPLPLIVGSSVG 1100
Db 1081 FFASTFTQVLTAAEIITDINPOLFVIEENAVTIPLIMKPTKAEVPT--GVIIGSIIA 1138
Qy 1101 GLLLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLLLAWTAGLWKLGFKKQYKDM 1163

RESULT 10
A33998
N;Integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 chain
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 21-Jul-2000
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: GB:X17033; NID:g33906; PIDN:CAA34894.1; PID:g33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
Biochem. J. 279, 419-425, 1991
R;Cacimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and GPIIc)
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:L24121; NID:g400342; PIDN:AAA16619.2; PID:g4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match
Best Local Similarity 26.9%; Pred. No. 1.3e-65;
Matches 329; Conservative 214; Mismatches 494; Indels 188; Gaps 43;

Qy 1 FNLDTENAMTFQ-ENARGFQSVQL---QGSRVVVGAPQEIIVAAQNRGLYQC--DYST 54
Db 30 YNVGLPEAKIFSGPSQFGYAVQQFINPKGNLLVGSFPMGPDVYKCPVDLST 89

```

```

Qy 55 GSCPIRLQ-----VPVEAVNMISGLSLAATTSPOLLACGPTVHTQCSNTYVKGLC 107
Db 90 ATCEKLNLTSTSIPIVNTKMTNMSLGLILTRNMGTFGLTCGLPLWAQCGNQTYYTTCVC 149
Qy 108 FLFGSNLRQOPQKFPFALRGCEQSDIAFLVDGSGSIIPHPRRRAKEPISVTMEOLK-- 165
Db 150 SDISPDF-QLSASFSPATQPCPSL-IDVVVVCDSESNISYPMW--AVKNFLEKPVQGLDIG 205
Qy 166 KSKITLFSLMQYSEEFRIHFTFEFQNNPNRSLIKPITQLLG-RHTATATGIRKVVRELFN 224
Db 206 PKTQVGLIYANNPRVFNLTNTYKKEEMIATSVTSQYGGDLTNFPAIQAIVARKVYS 265
Qy 225 ITNGARKNAFKILITDGEKFGDPLGYEDVPEADREGVIRYVIGV-----GDAFRSEK 279
Db 266 AASGGRRSATKVMVVVTDGESH-DGSMKAVIDQCNDHNLILRFGIAGVLGNALDNTKN 324
Qy 280 SRQELNTVASKPRPHVFOINNFEALKTIONQLREKIFAIEGTQTGSSSSFEHMEQEGF 339
Db 325 LIKIKAIASIPTEYFFNVSDAALLEKAGTLGEQIFSIEGTVQG-GDNFQEMSQVGF 383
Qy 340 SAAITSNGP--LLSTVGSYDMAGVFLYTSKEKSTFINMT--RVSDDMN-DAYLGVAAL 394
Db 384 SADYSSQNDILMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSSTLGSVAA 443
Qy 395 ILNRNVQSLVLAGPRYQHIGLVAMFRQNTGMWESNANV-----KGTQIGAYEGASLCSV 448
Db 444 ISTGESTHFVAGAPRANTGQIVLYSVN-----ENGNIIVIAHQRGDQIGSYFGSVLCSV 498
Qy 449 DVDSNGSTDLVLGAPHYEQTR--GGQVSVCPPLPRGQARWOCDAVLVGEQGPWRGFG 506
Db 499 DVDKDTITDVLVAGPAMYSDLKKEGRVYLFTIKKILGQHQ---FLEGPEGIENTRFG 555
Qy 507 AALTVLGDVNGDKLTDVAIGAPCEENRGAIVLPHGTSGSGISPSHSORIASG--KLSPR 564
Db 556 SAIALSDINMGFNDVIVGSPLENQNSGAVTIYNGHQT-IRTKYSQILGSDGAFRSH 614
Qy 565 LQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLLRSPQVLRVKAIMEFNPREVARNVFEEN 624
Db 615 LQYFGRLDGYDGLNGDSITDSIGAFQVQVQLWSQSIADVAIEASTPEKI--TLVNKN 672
Qy 625 DQVVKGEAGEVRVCLHVQKSTRDLREQIQSVVTYDIALD-----SGRPHSAVRVNETK 680
Db 673 AQII-----LKLCP---SAKFRPTKQNNQVAIVYNTILDADGFSRVRTSRGLPKENN 721
Qy 681 NSTRRQTVLGLTQTC--ETLKLQPNCTIEDPSPVLRNLNFSLVCTPLPSAFGLNPVLA 738
Db 722 ERCLQNMVYVNAQSCPEHIIYQEPS---DVNSLDLRVDISLENPGTS-----PALE 772
Qy 739 EDAQLRFTALFPFKNCGNDNICQDDLSITF-----SFMSLDCLVVGCPREFNVTVTRND 794
Db 773 AYSETAKVFSIPFKDCGEGDGLCISDLVDVRQIPAAQEQPFIVSQNKRLTFSVTLKNK 832
Qy 795 GEDSVTRQVTFPPPLDLSYRKVSTLQNRQSRWSRLACASAST-EVSGALKSTSCSINH 853
Db 833 RESAVNTGIVVDFSENLF-----ASFSLPVDGTEVTCOVAASQKSVACDVGY 880
Qy 854 PIFPENSEVTNITFDVDSKASIGNKLLKANVTSENMPRNTKTEFOLELPVKYAVYMW 913
Db 881 PALKREQQVTFITFDNLQ-NLQNASUSFQALSESQENKADNLVNLKIPILYDAEI- 938
Qy 914 VTSHGVSSTKYLNFNTASENTRVQHQYQVSNLQOR-----SLPISLPLV----- 958
Db 939 ---HLTRSTNINFIYSSDGNVPSIVHSFEDVGPKEIFSLKVTGTSVPVSMATVHIHQ 995
Qy 959 -----PVRLNQTIVWDRPQVTF--SENLSSTCHTKER 988
Db 996 YTKENKPLMYLTGVQTDKAGDISCNADINPLKIGQT---SSSVSPKSENFR---HTKE- 1047
Qy 989 LPSHSDFLAELKAPVNVCSIAVCORIQCDIPFFGIQEBFNATLKNLSFDWIKTSHNH 1048
Db 1048 -----LNCRTASCSNVTCLWKDVMHMKGBYFVNVTTRINWNGTFASFTQT 1091

```


Db 101 RKNPNTCEQLQSGSPGEGCGTKCLEERDNQWLGVTLSRQGENSGIVTCGHRWKNIFY 160
QY 96 TCSNTYVKGCLFGLFSGNLNQQPKFPEALRGCPQEDSDIAFLVDGSGSIIPDFRAXE 155
Db 161 MKSDNKLPTGICYWPSDLRLTSLK-----185
QY 156 FISTVMEQLKSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPITQLLGRTHATGI 215
Db 186 -----RMAPCYKDYT-----195
QY 216 RKVURELFNITNGARKNAFKLILITDGEKFGDPLGVEDVIPEADREGVIRYVIGVDAP 275
Db 196 -----RKFG-----200
QY 276 RSEKSRQELNTASKPRDRVFOINNPEALKTIQNLREKIFAIEGTQTGSSSFEHMS 335
Db 201 -----NFAS-----C 205
QY 336 QEGFSAAITNGPPLLSTVGYDNAGGVFLY--TSKEKSTFINMTRVDSMDNDAYLGYA-- 391
Db 206 QAGISSPYTQDLIVMGAPGSSYWTGTVFVNITTNQYKAFVD--RQNVKFGSYLGYSVG 263
QY 392 AAIIILNRVQSLVGLAPRYQHIGLVAMFRONTGMWESNAV---KGTQIGAYFGASLCS 447
Db 264 AGHFRSPHTTEVVGGAPQHGIQKAYIF-----SIDENELIVVEMKKGKLSYFGASVCA 319
QY 448 VDVSNGSTDLVLIGAPHYYEQTRGGQVSVCPLPRGORA--RMQCDVAVLXGEOQPGWRF 506
Db 320 VDLNADGFSDL-LVGAPMQSTIEBGRVFY--INSGMGAVMVNERVLVGSDDKYA-ARFG 376
QY 507 AALTUVLDVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHQRIAGSKLSPRLQ 566
Db 377 ESTANLGDIDNDGFEDIAIGAPQEDDLRGAVIYNGRV-DGISSTYSQRIEGQOISKSLR 435
QY 567 YFGQSLSGGDLTMDGLVDLTGCA--QGHVLLRSQPLVKAIMEFNPREVARNVPECN 624
Db 436 MFGQSSISGQIDADNGVYDVAVAFOSDSAVLRLTRPWIVEASLS-HPESVNRTRKPDCT 494
QY 625 DQVVKGEAGEVRVCLHVQKSTRDLRLEGQIQSVVTVDLALDSCR---PHSRVAFV--NET 679
Db 495 -----ENGLPSVCMHLTLCFSYKGEVPGYVLFVNVSLDVHRKAESPSRPFYFNGT 547
QY 680 KNSTRQTQVLGTQTCETLKLQLPNCIEDPSPVILRLNFSL-----VGTPLSAFNGLR 734
Db 548 SDVITGSIRVSSGKCRTHQAFMRKQVDRDILTPHVEATYHLGHVITKRNTEEPPLQ 607
QY 735 PVLAEDAQR-LFTALPFPEKNCNDNICQDDLSITFS-----FMSLDCLVGGCPREFNV 787
Db 608 PILQOKKEKDVIRKMINFARFCAYEN-CSADLQVSAKVGPLKPYENKTYLAVGSMKTI 666
QY 788 TVTVRNDGEDSYRTVTFPPPLDLSYRKVSTLQNRQSRSMWLACESASSTEVSGALKST 847
Db 667 NVSLFNAGDDAYETTLNVQLPTGLYFIKILDSEK-----QINCE-----VTESSGIVK-L 716
QY 848 SCSSINHPPIPFENSEVFNITFDVDSKASLGKLLKANVTSEN--NMPRTNKTEFOLELP 905
Db 717 ACSLGIYVDRLSRIDISFLDVSSLSRAHEDLSISVHASCENEGELDQVRNRTLTIP 776
QY 906 VKYAVVWVTSRGV--STKYLNFTASENTSRVNHQ-----YOVNLSGQRSPL-LSLVFL 957
Db 777 LRYEV--MLTVHGLVNPSTSVYSGSSENEPETCMAEKLNLTHFVINTGISMAPNVSVKIM 834
QY 958 VPVRLNQTIVWDRP--QVTFSENLSSTCHTKE-----RLPSHSDFLAELR 1000
Db 835 VP---NSFUPQDKLVNLDVQTTTQCHFKPHVGRECTFAQQGIAGTLTDIVKFLSKTD 891
QY 1001 KAPVWNCSTIA--VCQRIQDIPFFGQEEFNATLKGNSLPDWIKYKSHNLLIVSTAAIL 1058
Db 892 KR-LLYCMKADQHCLOFLQN---FGKVESG-----KEASVHIQLEGPSIL 933
QY 1059 FNDOSVTLIPGQAGFVRSQTEKVPFEPVNP-----1090
Db 934 EMDETSLSL-----KFEIKATAPPEPHPKVIELNKDENVAVHVFLEGLHHORPKRHF 983

QY 1091 -LPLIVSGVSGGLLIALITALYALYKLGFFKQVQKMMSE 1128
Db 984 TIIITITISLLGLIVLLISCVWMAKAGFFKQVKSILQE 1022

RESULT 13

S06046

Integrin alpha-4 chain precursor - human

N:Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C:Species: Homo sapiens (man)
C>Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 24-Sep-1999
R:Accession: S06046; A39355; D28018
R:Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989

A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other int
A:Reference number: S06046; MUID:89356603; PMID:2788572
A:Accession: S06046

A:Molecule type: mRNA

A:Residues: 1-1038 <TA>

A:Cross-references: GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1; PID:g33946
R:Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A:Title: Characterization of the alpha-4 integrin gene promoter.
A:Reference number: A39355; MUID:91239513; PMID:2034655
A:Accession: A39355

A:Molecule type: DNA

A:Residues: 1-93 <ROS>

A:Cross-references: GB:M62841

R:Takada, Y.; Strominger, J.L.; Hemler, M.E.

Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987

A:Title: The very late antigen family of heterodimers is part of a superfamily of mole

A:Reference number: A94151; MUID:87204112; PMID:3033641

A:Accession: D28018

A:Molecule type: protein

A:Residues: 40-50,'E',52-53 <TA2>

C:Genetics:

A:Gene: GDB:ITGA4; CD49D

A:Cross-references: GDB:128032; OMIM:192975

A:Map position: 2q31-2q32

C:Superfamily: integrin alpha-4 chain

C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; trans
F:1-39/Domain: signal sequence #status predicted <SIG>
F:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match

Best Local Similarity 10.8%; Score 635; DB 2; Length 1038;

Matches 249; Conservative 154; Mismatches 370; Indels 212; Gaps 39;

QY 272 GDAFRSEKSRQELNTVASKPPRD-----HVFOINNPEALKT-----IQN 310
Db 121 GKTCLLEERDNQWLGVTLSRQGENSGIVTCGHRWKNIFYIKENKLP TGCGYCVPPDLRT 180
QY 311 QUREXI-----FAIEGTQTGSSSFEHMSQEGFSAAITNGSPLLTGVSY 356
Db 181 ELKSLIAPCYQDYVKKFGENFA-----SCQAGISSPYTKDLIVMGAPGS 225
QY 357 DNAGGVFLY---TSKEKSTFINMTRVDSMDNDAYLGYA--AAIILNRVQSLVGLAPRQ 411
Db 226 YWTGSLFVNITTNKYKAFDRQNVKF---GSYLGSYVAGHFRSQHTTEVVGGAPQHE 282
QY 412 HIGLVAMFRONTGMWESNAV---KGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYY 467
Db 283 QIGKAYIF-----SIDEKELNIHEMKKGLSYFGASCAVDLNDAGFSDL-LVGAPMQS 337
QY 468 EQTRGQVSVCLPRGQRARWQC-DALVYGEQPGWRFGAALTVLVDGYNQKLTDAVIG 526
Db 338 TIREGRVFY--INSGGAVVMAMETNLVGSDDKYA-ARFGESIVNLGIDINDGFEDVAG 395
QY 527 AGCEDNRGAVYLFHGTSGSGISPSHQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDL 586
Db 396 APQEDDLQAGAIYINGRA-DGISSTFSQRIEGLQISKLSLMSFGQSTISGQIDADNNGYVDV 454

A;Description: Alpha SU2, a sea urchin integrin which binds laminin.

A;Reference number: Z21035

A;Accession: T31437

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1041 <HER>

A;Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1

A;Experimental source: developmental stage embryo

C;Function:

C;Superfamily: binds laminin

C;Superfamily: integrin alpha-2b chain

Query Match

Best Local Similarity 10.5%; Score 614.5; DB 2; Length 1041;

Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

336 QEGFSAITNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS-----DWDNAYLG 389

181 QAGFSGIIFSDNSALVNGAFSGYYLQGIYVQSLNRSV-VQATQESNTGTSYFDSNRYG 239

390 YAAAI--ILNRVQSLVLAGAPRYQHI--GLVAMFRQNTGMWESNANVKGTOIGAYFGASLC 446

240 YSLALGDFNGGVQDYVVGTPRAESLMLVAIFDQNLNQFN---QWGTQIVAYFGYSVT 296

447 SVVDNSGSTDLVLIGAPHYYEOTRGQSVSCPLPRQARWQCDVLYGEQ----- 498

297 VVDI--NNDTVDDLVLGAPMYMDGPAIQ-----RWEAGAVVYVLQNPVGPQA 343

499 -----GQPCWRGGAALTVDVNGDKLTDVAICAPGEEDNRGAVYLPFGTSGS 546

344 SNRLSLSSTLIIGQIRSRPGLSTIASIGDSNQDGFNDVAIGAPYEGDDAGAVIYHG--SAN 402

547 GISPSHSQRIAGSKLS--PRLOYFGQSLSGGQDITMDGLVLTVGAQ--GHVLLRSQPV 603

403 GLKSTPAQVLTPTSLHSGITTFGFSLQGGQDMKXKYPDLLVGAESANTAVLIRTPVV 462

604 RVKAIMFENPREVARNVFENQDVVKGEAGEVRVCLHVQKSTRDLRREGQIQSV----- 658

463 SLDATLNTPE-----IGINLENTKYE--LADGTMVTSFIAMT 497

659 -----VTYDLALDSG--RPHSRVAFNETKNSTRQTVLGL--TQTCETLKLQ 702

498 CFTYTGNYLPHDIDISYTVTVDSGIIANRRAMFVNDMSIITKTRRLAVSTQFCPLRAY 557

703 LPNCIEDPVSPVILRLNFSLVGTPLSAFGN-----LRPVLAEDAQRLEFALPPEK 753

558 VGNSEDKLTPIKVTLOYDL-----NDESRLOPHEILPIIDMATMSTQTKQVSIQN 609

754 NCGNDNICODDLSITFSFMSLCLVGGPREFNVTVTRNDGEDSYRTQVTFPPPLDLSY 813

610 NCVN-NICIPDLQVTVT--PNLPNIVIGTQELTLDVSLNRRGEDAFQSSLSVYVPLQLQF 667

814 RKVSTLQNRQSRWRACESASSTVSGALKSTCSINHIPIFEN-----SEVTFNIT 867

668 VRL-----ERKANMDFSVTCSSESD-----LRIITCDTGNPMVGKNILEFGLTLSTFQVS 717

868 FDVDS-----KASLGKLLKKNVTSNNMPTNKTEFQLELPKYAVYVMTVSHGVSTK 922

718 GDXDSIEFYFAKESNS--EDPNTLENLNMVTVTVCTCLLKSASYPEIVMYSTQED 775

923 YL-----NFTASENTSRVMOHQYQV-----SNLQORSIPIS-----LVFLVPV 960

776 YVVPVPPFAKASEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPKQKDEGEYLFYLLGI 834

961 RLNQTVIWRPQ-----VTFSENLSSTCHTKERLPSHSDFLAELRKAPVV 1005

835 MTEBEGVTCQLTQCKANPEGVKLEPSTKAKLSNSTTQVSGRRREPEVAEALQTDN--VI 892

1006 NCSIAVCQRIQCDDIPFGIOEEFNAT-----LKGNLSPDWYIKTSHNHLIIYSTAEILF 1059

893 YCASDSCVLINCTI-----DEINASKSVVRIILGRF--W-----ERTF 928

1060 NDSVFTLLPGGAGFVRQSTETKVE--PFEVFPNP-----LP----- 1092

Db 929 QKAVSELTPVQQTATIASASAAVKTIPTYNIPLPDPSDSTKASTLVTTELVPPVPTPIAW 988

QY 1093 -LIVGSSVGGLLLLALITAAALYKLGFFKR-----QYKDMMS--EGGPP 1132

Db 989 WIIVSVLGGIILLIILGLWKCGFFERKKPGEKEYAPVASADKGGPP 1038

Search completed: November 25, 2003, 14:22:04

Job time : 18.9579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.17418 Seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481b-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTFQENARFGQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5846.5	99.5	1152	1	ITAM_HUMAN
2	4470	76.1	1153	1	ITAM_MOUSE
3	3469	59.0	1163	1	ITAX_HUMAN
4	3411	58.0	1162	1	ITAD_HUMAN
5	1557.5	26.5	1170	1	ITAL_HUMAN
6	1538.5	26.2	1163	1	ITAM_MOUSE
7	1161.5	19.8	1167	1	ITAE_MOUSE
8	1153	19.6	1179	1	ITAH_HUMAN
9	1103.5	18.8	1151	1	ITAI_HUMAN
10	1093.5	18.6	1189	1	ITAH_HUMAN
11	1085	18.5	1170	1	ITAT_BOVIN
12	1071	18.2	1178	1	ITAT_MOUSE
13	1068	18.2	1181	1	ITAT_HUMAN
14	1064.5	18.1	1167	1	ITAG_HUMAN
15	1062	18.1	1180	1	ITAI_RAT
16	663	11.3	1039	1	ITAT_MOUSE
17	635	10.8	1038	1	ITAT_HUMAN
18	630	10.7	1035	1	ITAT_HUMAN
19	593.5	10.1	1032	1	ITAT_XENLA
20	571.5	9.7	1066	1	ITAT_CRISP
21	567.5	9.7	1053	1	ITAT_MOUSE
22	555.5	9.5	1053	1	ITAT_MOUSE
23	545.5	9.3	1050	1	ITAT_XENLA
24	542.5	9.2	1034	1	ITAT_CHICK
25	537.5	9.1	1130	1	ITAT_HUMAN
26	535	9.1	1044	1	ITAT_MOUSE
27	532.5	9.1	1066	1	ITAT_HUMAN
28	532	9.1	1049	1	ITAT_HUMAN
29	531.5	9.0	1072	1	ITAT_CHICK
30	526	9.0	1048	1	ITAT_HUMAN
31	517	8.8	1044	1	ITAT_CHICK
32	512.5	8.7	1091	1	ITAT_MOUSE
33	498	8.5	1179	1	ITAT_MOUSE

ALIGNMENTS

RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Fahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	496	8.4	1396	1	ITAT_DROME
35	491.5	8.4	1146	1	ITAI_DROME
36	491	8.4	1033	1	ITAB_MOUSE
37	489.5	8.3	1039	1	ITAB_HUMAN
38	489	8.3	1025	1	ITAB_HUMAN
39	486	8.3	126	1	ITAM_CAVPO
40	476	8.1	1181	1	ITAT_HUMAN
41	470	8.0	1106	1	ITAT_RAT
42	467	7.9	1226	1	PAT2_CABEL
43	445.5	7.6	1139	1	INAI_CAEEL
44	424	7.2	1115	1	ITAT_DROME
45	385	6.6	1000	1	ITAT_DROME

P12080	drosophila
O24247	drosophila
O9qum0	mus musculus
P08514	homo sapien
P53708	homo sapien
P11578	cavia porce
Q13683	homo sapien
Q63258	rattus norv
P34446	caenorhabdi
O03600	caenorhabdi
O44386	drosophila
O9wim8	drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7].
RP SEQUENCE OF 1-9 FROM N.A.
RX TISSUE=Blood;
RC MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8].
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol;
RT conservation across species and homology to platelet IIB/IIIA.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9].
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10].
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11].
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,
RA Horton N.C., Kelley L.L., Wildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment.";
RL Structure 6:923-935(1998).
RN [12].
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
CC OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
CC ASSOCIATES WITH BETA-2.
CC -!- SURCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide cd11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S5227; AAB24821.1; -
DR EMBL; S52152; AAB24821.1; JOINED.
DR EMBL; S52153; AAB24821.1; JOINED.
DR EMBL; S52154; AAB24821.1; JOINED.
DR EMBL; S52155; AAB24821.1; JOINED.
DR EMBL; S52157; AAB24821.1; JOINED.
DR EMBL; S52159; AAB24821.1; JOINED.
DR EMBL; S52161; AAB24821.1; JOINED.
DR EMBL; S52164; AAB24821.1; JOINED.
DR EMBL; S52165; AAB24821.1; JOINED.
DR EMBL; S52167; AAB24821.1; JOINED.
DR EMBL; S52169; AAB24821.1; JOINED.
DR EMBL; S52170; AAB24821.1; JOINED.
DR EMBL; S52173; AAB24821.1; JOINED.
DR EMBL; S52174; AAB24821.1; JOINED.
DR EMBL; S52180; AAB24821.1; JOINED.
DR EMBL; S52181; AAB24821.1; JOINED.
DR EMBL; S52184; AAB24821.1; JOINED.
DR EMBL; S52189; AAB24821.1; JOINED.
DR EMBL; S52191; AAB24821.1; JOINED.
DR EMBL; S52192; AAB24821.1; JOINED.
DR EMBL; S52203; AAB24821.1; JOINED.
DR EMBL; S52212; AAB24821.1; JOINED.
DR EMBL; S52213; AAB24821.1; JOINED.
DR EMBL; S52216; AAB24821.1; JOINED.
DR EMBL; S52219; AAB24821.1; JOINED.
DR EMBL; S52220; AAB24821.1; JOINED.
DR EMBL; S52221; AAB24821.1; JOINED.
DR EMBL; S52222; AAB24821.1; JOINED.
DR EMBL; S52226; AAB24821.1; JOINED.
DR EMBL; M76724; AAA58410.1; -
DR EMBL; M84477; AAA51960.1; -
DR PIR; A31108; RMHUIB.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01813; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.

Query Match		99.5%;	Score 5846.5;	DB 1;	Length 1152;		
Best Local Similarity		99.2%;	Pred. No. 0;				
Matches 1128;		Conservative	7;	Mismatches	1;	Indels	1; Gaps 1;
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCBPI	60				
DB	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCBPI	76				
QY	61	RLQVPVAVNMSLGLSLAATTSPOLLACGPTVHQCSTENTYVKGFLCFLGSLNLRQPOK	120				
DB	77	RLQVPVAVNMSLGLSLAATTSPOLLACGPTVHQCSTENTYVKGFLCFLGSLNLRQPOK	136				
QY	121	FPEALRGCPQEDSDIAFLVDGSGIIPHPDFRRAKEFISTVMEQLKSKTFLSLMQVSEEF	180				
DB	137	FPEALRGCPQEDSDIAFLVDGSGIIPHPDFRRAKEFISTVMEQLKSKTFLSLMQVSEEF	196				
QY	181	RIHFTPKFONNPNRSLKIPITQLLGRTHATGIRKVVRELNIINGARKNAFKILILI	240				
DB	197	RIHFTPKFONNPNRSLKIPITQLLGRTHATGIRKVVRELNIINGARKNAFKILIVI	256				
QY	241	TDGEKFGDPLGYEDVLPADREGVIRVVGVDGAPFRSEKSRQELNTVASKPPEDHVFQIN	300				
DB	257	TDGEKFGDPLGYEDVLPADREGVIRVVGVDGAPFRSEKSRQELNTVASKPPEDHVFQIN	316				
QY	301	NFEALXTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATITNSGPLLSTVGSYDWAG	360				
DB	317	NFEALXTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATITNSGPLLSTVGSYDWAG	376				
QY	361	GVFLYTSKEKSTFINTRVDSNDNDAYLVYAAAIILNRVQSLVLGAPRYQHIGLVAMFR	420				
DB	377	GVFLYTSKEKSTFINTRVDSNDNDAYLVYAAAIILNRVQSLVLGAPRYQHIGLVAMFR	436				
QY	421	QNTGMESNANVKGTOIGAYFGASLCSDVDNSGSDTLVLIGAPHYYEQTGRGQSVCP	480				
DB	437	QNTGMESNANVKGTOIGAYFGASLCSDVDNSGSDTLVLIGAPHYYEQTGRGQSVCP	496				
QY	481	PRQORARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	540				
DB	497	PRG-RARWQCDVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	555				
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLRLRSQ	600				
DB	556	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLRLRSQ	615				
QY	601	PVLVRKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLVHVKQSTRDLREGQIOSVVT	660				
DB	616	PVLVRKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLVHVKQSTRDLREGQIOSVVT	675				
QY	661	YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF	720				
DB	676	YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF	735				
QY	721	SLVGTPLSAPGNLRPVLAEADAQLFTLFPFKNCGNDNICQDDLSITFSFMSLDCLVWG	780				
DB	736	SLVGTPLSAPGNLRPVLAEADAQLFTLFPFKNCGNDNICQDDLSITFSFMSLDCLVWG	795				
QY	781	GPREFNVTVVRNDGEDSYQTQVTFPPDLVSRKYSTLQNRQSRWRLACESASSTEV	840				
DB	796	GPREFNVTVVRNDGEDSYQTQVTFPPDLVSRKYSTLQNRQSRWRLACESASSTEV	855				
QY	841	SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF	900				
DB	856	SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF	915				
QY	901	QLELPVKYAYVMVYTSHGVTSKYLNFTASENTSRVMQHQYQVSNLQORSLPISLVLVPV	960				
DB	916	QLELPVKYAYVMVYTSHGVTSKYLNFTASENTSRVMQHQYQVSNLQORSLPISLVLVPV	975				
QY	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIACVORIQCIP	1020				
DB	976	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIACVORIQCIP	1035				

1021 PFGIOEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPQOGAFVRSQTET 1080
1036 PFGIOEFNATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPQOGAFVRSQTET 1095
1081 KVEPEFVNPLPIVGVSSVGGLLLALITAAALYKLGFFKRYKDMSEGGPGCAEPQ 1137
1096 KVEPEFVNPLPIVGVSSVGGLLLALITAAALYKLGFFKRYKDMSEGGPGCAEPQ 1152

RESULT 2
ITAM_MOUSE
ID ITAM_MOUSE STANDARD; PRT: 1153 AA.
AC P05555;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MO1). ITGAM.
GN ITGAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88312584; PubMed=3044779;
RA Pytela R.;
RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";
RL EMBO J. 7:1371-1378(1988).
RN [2]
RP SEQUENCE OF 11-45 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=86287312; PubMed=2942940;
RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;
RT "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
RN [3]
RP SEQUENCE OF 17-28.
RX MEDLINE=85188276; PubMed=3887182;
RA Springer T.A., Teplow D.B., Dreyer W.J.;
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MOST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLomerulonephritis. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X07640; CAA30479.1; -;
EMBL; M14293; AAA39484.1; -;
PIR; S00551; S00551.
HSSP; P11215; 1ABX.

MGI; 96607; Itgam.

InterPro; IPR00413; Integrin_alpha.

InterPro; IPR002035; VWF_A.

Pfam; PF01839; FG-GAP; 3.

Pfam; PF00357; Integrin_A; 1.

Pfam; PF00092; vwa; 1.

PRINTS; PR01185; INTEGRINA.

PRINTS; PR00453; VWFADOMAIN.

SMART; SM00191; Int_alpha; 5.

SMART; SM00327; WVA; 1.

PROSITE; PS00242; INTEGRIN_ALPHA; 1.

PROSITE; PS02034; VWF_A; 1.

Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

Signal; Calcium; Repeat.

SIGNAL 1 16

CHAIN 17 1153

DOMAIN 17 1105

TRANSMEM 1106 1129

DOMAIN 1130 1153

REPEAT 31 84

REPEAT ? ?

DOMAIN 164 350

REPEAT 337 400

REPEAT 401 452

REPEAT 454 515

REPEAT 517 575

REPEAT 580 632

CA_BIND 465 473

CA_BIND 529 537

CA_BIND 592 600

SITE 1132 1136

DISULFID 66 73

DISULFID 105 123

DISULFID 654 711

DISULFID 770 776

DISULFID 999 1023

DISULFID 1028 1033

CARBOHYD 58 58

CARBOHYD 86 86

CARBOHYD 391 391

CARBOHYD 696 696

CARBOHYD 734 734

CARBOHYD 772 772

CARBOHYD 801 801

CARBOHYD 881 881

CARBOHYD 907 907

CARBOHYD 941 941

CARBOHYD 980 980

CARBOHYD 994 994

CARBOHYD 1022 1022

CARBOHYD 1045 1045

CARBOHYD 1051 1051

CARBOHYD 1076 1076

SEQUENCE 1153 AA; 127480 MW; 178DB988AECB0343 CRC64;

Query Match 76.1%; Score 4470; DB 1; Length 1153;

Best Local Similarity 73.9%; Pred. No. 4.5e-295;

Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;

1 FNLDTEHMTFOENAKGFCQVNVQJGGTSVVVAAPQEAQKAVNQTGALYCCDYSTRCHPI 76

17 FNLDTEHMTFOENAKGFCQVNVQJGGTSVVVAAPQEAQKAVNQTGALYCCDYSTRCHPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHOTCSENTYVKGCLFLFGLNLRQOPQK 120
DB 77 PLQVPEAVNMSLGLSLAVSTVPQQLACGPTVHOTCSENTYVKGCLFLFGLNLRQOPQK 136
QY 121 FPEALRGCEQSDSDIAFLVDGSGSIIPHDPRAKFPISTWMBQLKSKTFLPSLMQYSEBF 180
DB 137 FPEALRGCEQSDSDIAFLVDGSGSIIPHDPRAKFPISTWMBQLKSKTFLPSLMQYSEBF 196
QY 181 RIHFTPEKFNPNPRLSKPITOLLGRTHRTATGRKVVRELFININGARKNAFKLIL 240
DB 197 RIHFTFNDPKRNPSPRSHVSPIKQLNGRTAGIRKVVRELFININGARKNAFKLIL 256
QY 241 TDGEKFGDPLGYEDVPIPEADRGVIRYVIGVGNFKNPQSRRELDITASKPAGEHVFOVD 300
DB 257 TDGEKFGDPLGYEDVPIPEADRGVIRYVIGVGNFKNPQSRRELDITASKPAGEHVFOVD 316
QY 301 NFEALNTIQNLOREKIFAIEGTQTGSSSEFHEHMSQEGFSAITSNGPLLSVGVSDWAG 360
DB 317 NFEALNTIQNLOREKIFAIEGTQTGSSSEFHEHMSQEGFSAITSNGPLLSVGVSDWAG 376
QY 361 GYPLTYSKEKSTFINNTRVDSMDNDAYLGYAAAILNRVOSILVLCAPRYOHIGLVAMPR 420
DB 377 GAFLYTSKOKVTPINTTRVDSMDNDAYLGYAAAILNRVOSILVLCAPRYOHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSCPL 480
DB 437 ENGTWEPHTSINKSGIYSYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSCPL 496
QY 481 PRQARWQCDVLYGEGQOPWGRFGAALTGLGVNGDKLTJDAI GAPGEDNRGAYVLP 540
DB 497 PRG-RARWQCEALLHGDQHPWGRFGAALTGLGVNGDKLTJDAI GAPGEDNRGAYVLP 555
QY 541 HGTSGSGISPSHSORITAGSKLSPLOYFGOSLGGDLTMDGLVDITVGAQGHVLLRSQ 600
DB 556 YGASIASLSASHSHRIIGAHFSPGLQYFGOSLGGDLTMDGLVDITVGAQGHVLLRSQ 615
QY 601 PVLVRKAIMEFNPREVARNVFECDQVWKGKEAGEVRVCLVHVKSTRDRREGQIQSVVT 660
DB 616 PVLRLTEATHEFSPKKVARSVFACQEVKKNKDAAGEVRVCLVHVKSTRDRREGQIQSVVT 675
QY 661 YDLALDGRPHSRVAFNETKNSRRTOTVGLTQTCETLKLQIPNCIEDBPVPIVRLNF 720
DB 676 YDLALDGRPHSRVAFNETKNSRRTOTVGLTQTCETLKLQIPNCIEDBPVPIVRLNF 735
QY 721 SLVGTPLSAGNLRPVLAEDAORLFTALPFPEKNCNGNDICQDDLSITFSFMSLDCVLVG 780
DB 736 TLVGEPLRSFNGNLRPVLAEDAORLFTALPFPEKNCNGNDICQDDLSITFSFMSLDCVLVG 795
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRYKVKVSTLQNSQRSWRL-ACESASSTE 839
DB 796 GPQDFNMSVTLRNDGEDSYRTQVTFPPDLVSRYKVKVSTLQNSQRSWRL-ACESASSTE 855
QY 840 VSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKYTSNNPRTNKT 899
DB 856 GHGALKSTTWNHPIPPANSEVTFNITFDVDSKASLGNKLLKANKYTSNNPRTNKT 915
QY 900 FQLELPKYAVYVWTSVSHGVSTKYVNFASNTSRVWQHYOVSNLQORSPLISLVLPV 959
DB 916 FQLELPKYAVYVWTSVSHGVSTKYVNFASNTSRVWQHYOVSNLQORSPLISLVLPV 975
QY 960 VRLNQTVIDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORICDI 1019
DB 976 VQINNVTVWHPQVIFSONLSSACHTEQKSPHNSFRDQLERTPVLNCSVAVCKRIQCDL 1035
QY 1020 PFFGQIEFNATLGNLSFDWYIKTNSHLLIIVSTAEILFNDSVFTLLPGQAFVRSOTE 1079
DB 1036 PSFNTQIEFNATLGNLSFDWYIKTNSHLLIIVSTAEILFNDSVFTLLPGQAFVRSOTE 1095
QY 1080 TKVEPPEVNPDLPLIVGSSVGGLLLLALITAAALKLGFKKQYKQDMSEGGPPQAEQ 1137
DB 1096 TKVEPPEVNPDLPLIVGSSVGGLLLLALITAAALKLGFKKQYKQDMSEGGPPQAEQ 1153

RESULT 3
 ID ITAX HUMAN STANDARD; PRT; 1163 AA.
 AC P20702;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95
 DE alpha chain) (leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
 GN ITGAX OR CD11C.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88166645; PubMed=3327687;
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
 RT "cDNA cloning and complete primary structure of the alpha subunit of
 RT a leukocyte adhesion glycoprotein, p150,95";
 RL EMBO J. 6:4023-4028(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90153906; PubMed=2303426;
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RT "Genomic structure of an integrin alpha subunit, the leukocyte
 RT p150,95 molecule";
 RL J. Biol. Chem. 265:2782-2788(1990).
 RN [3]
 RP ERRATUM.
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RL J. Biol. Chem. 265:12750-12751(1990).
 RN [4]
 RP SEQUENCE OF 20-43.
 RX MEDLINE=87167596; PubMed=3549901;
 RA Miller L.J., Wiebe M., Springer T.A.;
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1
 RT and p150,95 leukocyte adhesion proteins";
 RL J. Immunol. 138:2381-2383(1987).
 CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
 CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
 CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
 CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
 CC ASSOCIATES WITH BETA-2
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
 CC GRANULOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to licenses@isb-sib.ch).
 CC -----
 DR EMBL; M81695; AAA59180.1; -;
 DR EMBL; Y00093; CAA68283.1; -;
 DR EMBL; M29165; -; NOT ANNOTATED CDS.
 DR EMBL; M29487; AAA51620.1; ALT SEQ.
 DR EMBL; M29482; AAA51620.1; JOINED.
 DR EMBL; M29483; AAA51620.1; JOINED.
 DR EMBL; M29484; AAA51620.1; JOINED.
 DR EMBL; M29485; AAA51620.1; JOINED.

DR EMBL; M29486; AAA51620.1; JOINED.
 DR PIR; A36584; RWHUIC.
 DR PDB; IN3Y; 18-FEB-03.
 DR Genew; HGNC:6152; ITGAX.
 DR MIM; 151510; -;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 DR GO; GO:0004872; F:receptor activity; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWFA.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00327; VWA_1.
 DR SMART; SM00191; Int_alpha; 5.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 1163
 FT DOMAIN 20 1107
 FT TRANSMEM 1108 1128
 FT DOMAIN 1129 1163
 FT REPEAT 34 87
 FT REPEAT ? ?
 FT REPEAT 165 351
 FT DOMAIN ? ?
 FT REPEAT 402 453
 FT REPEAT 455 517
 FT REPEAT 518 576
 FT REPEAT 581 633
 FT CA_BIND 466 474
 FT CA_BIND 530 538
 FT CA_BIND 593 601
 FT SITE 1131 1135
 FT DISULFID 69 76
 FT DISULFID 108 126
 FT DISULFID 655 712
 FT DISULFID 771 777
 FT DISULFID 848 863
 FT DISULFID 998 1022
 FT DISULFID 1027 1032
 FT CARBOHYD 61 61
 FT CARBOHYD 89 89
 FT CARBOHYD 392 392
 FT CARBOHYD 697 697
 FT CARBOHYD 735 735
 FT CARBOHYD 899 899
 FT CARBOHYD 939 939
 FT CARBOHYD 1050 1050
 FT CONFLICT 490 490
 FT CONFLICT 756 756
 FT CONFLICT 1163 AA; 127885 MW; 6C4E19C3P62A473 CRC64;
 SQ SEQUENCE

Query Match 59.0%; Score 3469; DB 1; Length 1163;
 Best Local Similarity 61.0%; Pred. No. 3.8e-227;
 Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;
 QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVGAPOEIVAAANQGSLYOCDYSTGSCBPI 60
 DB 20 FNLDTEELTAFRVDSAGFGSDSVVQVANSVWVGAPQKITAANQTGLVQCGYSTGACBPI 79
 QY 61 RLQVPEAVNNMGLSLGSLAATTSPPQLACGPTVHOTCSNTYVKGCLFGLFNSLRQOPK 120
 DB 80 GLQVPEAVNNMGLSLGSLASTTSPSOLLACGPTVHHECGRNMYLTGLCLLGLPT--QLTOR 137
 QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKKSTLFSLMQYSEF 180
 DB 138 LPVSRQECPCROEQDIVFLIDGSGSSSRNFATMNFVRAVISQFORPSTQFSLMQFSNKF 197

DR EMBL; U40274; AAB60634.1; --
DR EMBL; U40275; AAB60635.1; --
DR EMBL; U40276; AAB60636.1; --
DR EMBL; U40277; AAB60637.1; --
DR EMBL; U40278; AAB60638.1; --
DR EMBL; U40279; AAB60639.1; --
DR EMBL; U40280; AAB60640.1; --
DR EMBL; AF187881; AAF62875.1; --
DR HSP; P11215; IABX.
DR Genew; HGNC:6146; ITGAD.
DR MIM; 602453; --
DR GO; 0008305; C: Integrin complex; TAS.
DR GO; 0004895; P: cell adhesion receptor activity; TAS.
DR GO; 0016337; P: cell-cell adhesion; NAS.
DR GO; 0007160; P: cell-matrix adhesion; NAS.
DR GO; 0006955; P: immune response; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00327; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02234; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium;
KW Magnesium.
FT SIGNAL 1 17
FT CHAIN 18 1162
FT DOMAIN 18 1100
FT TRANSMEM 1101 1124
FT DOMAIN 1125 1162
FT REPEAT 32 85
FT REPEAT ?
FT DOMAIN 150 332
FT REPEAT 350 400
FT REPEAT 401 452
FT REPEAT 454 516
FT REPEAT 518 576
FT REPEAT 581 633
FT CA_BIND 465 473
FT CA_BIND 530 538
FT CA_BIND 593 601
FT SITE 1127 1131
FT DISULFID 67 74
FT DISULFID 106 124
FT DISULFID 655 710
FT DISULFID 769 775
FT DISULFID 846 861
FT DISULFID 994 1018
FT DISULFID 1023 1028
FT CARBOHYD 59 59
FT CARBOHYD 87 87
FT CARBOHYD 99 99
FT CARBOHYD 391 391
FT CARBOHYD 691 691
FT CARBOHYD 733 733
FT CARBOHYD 873 873
FT CARBOHYD 957 957
FT CARBOHYD 1046 1046
FT CONFLICT 500 500
FT CONFLICT 515 518
FT CONFLICT 825 825
FT CONFLICT 984 984
SQ SEQUENCE 1162 AA; 126885 MW; F296A1A35455D77D CRC64;

Query Match
Best Local Similarity 59.4%; Pred. No. 3.3e-223;
Matches : 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVYVVGAPQEIIVAAQNRGSLYQCDSYTGSCPEI 60

DB 18 FNLDVEEPTIFQEDAGGFGQSVVQLOGSRVYVVGAPLEVVAAQNRGLYDCAATGMCQPI 77
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLPGSNLRQPOPK 120
DB 78 PLHIRPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGICFLPGSNLRQPOPK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKKSKTFLSLMQVSEEF 180
DB 137 VPDATPECPHQEMDIVFLDGGSGSDQNDFNQMGFVQAVMGQFEGCTDTLFFALMQVSNLL 196
QY 181 RIHFTPEKEPQNNPNRSLIKPIITOLLGRTHATATGRKVVRELPNITNGARKNAFKILLI 240
DB 197 KIHFTFTQRTSPSQSLVDPIVLQKGLTFTATGILTVVTQLFHHKNGARKSAKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDFRSEKSRQELNVTASKPRDHDVQJN 300
DB 257 TDGQYKDPLEYSVDIPQAEKAGIIRYALGVGHAFCGPTARQELNVTISSAPQDHFVKVD 316
QY 301 NFEALKTIQNLREKIIFAIETGTGSSSSFEHMSQEGFSAATITNGPILLSTVGSYDWAG 360
DB 317 NFAALGSIQQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGFSWSG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAAILNRVQSLVGLGAPRYQHIGLVAMPR 420
DB 377 GAFLYPPNMSPTFINNSQENVMDRDSYLGSTELALWKGQVNLVGLGAPRYQHTKAVFT 436
QY 421 QNTGHWESNANVKGTCIGAYFGASLCSDVDVDSNGSTDVLIGAPHYVETRGQGVSCPL 480
DB 437 QVSRQWRKKAETVGTQIGSYFGASLCSDVDVDSNGSTDVLIGAPHYVETRGQGVSCPL 496
QY 481 PRGQARMQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
DB 497 PRGQVQWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFCQSLSGGGDLTMGDLVDTVGAQGHVLLRSQ 600
DB 557 HGASESGISPSHSQRIAGSKLSPRLQYFCQSLSGGGDLTMGDLVDTVGAQGHVLLRSQ 616
QY 601 PVLRYKATIMEFNPREVARNVFECDQVVKGEVUCLHVQKSDRDLREGQISVVT 660
DB 617 PVLKVGVMRFPVEVAKAVYRCWEEKPSALEAGDATVCLTQKSLDQL--GDIQSSVR 674
QY 661 YDLALDSGRPHSRVFNETKSTRQTQVLGTQTCETLKQLPNCIEDPVSPIVLRNF 720
DB 675 FDLALDPCRLTSRAIFNETKNTLTERTGLGHCETLKLLJLPCDVVDVSPILHLNF 734
QY 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPPFKNCGNDNICDDLSITFSWLSCLVVG 780
DB 735 SLVREPIPSQNLRLPVLAVGSDQLFTASLPFEKNCQDGLCEGLGVTLSFSLQTLTVG 794
QY 781 GPREENVTYVENDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSWPLACESASSTE 840
DB 795 SLELNVIYVWNAGEDSYGTVVSUYPAGLSHRRVSGAQKQPHOSALELACETV--PTED 853
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANKYTSNNMPTWKTFP 900
DB 854 EG-LSSRCSVNHPIFHEGNGCTFIVTDFVSKATLGDRLMRASSENKASSKATF 912
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNF--TASENTSRYMHQYQVSNLQGRSLPISLVLPV 959
DB 913 QLELPVKYAVYVMTVSHGVSTKYLNF--TASENTSRYMHQYQVSNLQGRSLPISLVLPV 972
QY 960 VPLNQTVIWRDQVTFSENLSSTCHTKERLPHSDFLAELRAKPVVNCVNCIAVCQRQCDI 1019
DB 973 VLLNGVAVWVWMEAPSQSL--PCVSEKPPQHSDFLTQISPSMLDCCDIADCLQPCDV 1030
QY 1020 PFFGQIEEFNATLKNLSDFDWIKYKSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTE 1079
DB 1031 PFSVQIEELDFTLKNLSFGWVRETLOKXVLVSVVAEITFTDTSVSVQLPGQCAFVRSOTE 1090
QY 1080 TKVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKRYQKMMSE 1128

Db 1091 MVLBEDEVYNAIPIMSSV GALLLALITATLYKLGLFKRHYKEMLED 1139

RESULT 5

ITAL HUMAN STANDARD: PRT; 1170 AA.

AC P20701; Q43746;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).

DE ITGAL OR CD11A.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.

RX MEDLINE=89139587; PubMed=3537322;

RA Larson R.S., Corbi A.L., Berman L., Springer T.;

RT "Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily.";

RL J. Cell Biol. 108:703-712 (1989).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RX MEDLINE=99425270; PubMed=10493829;

RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;

RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";

RL Genomics 60:295-308 (1999).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.

RX MEDLINE=96036067; PubMed=7479767;

RA Qu A., Leahy D.J.;

RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin.";

RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281 (1995).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.

RX MEDLINE=96398682; PubMed=8805579;

RA Qu A., Leahy D.J.;

RT "The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin.";

RL Structure 4:931-942 (1996).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.

RX MEDLINE=99425288; PubMed=10493852;

RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G., Cottens S., Weitz-Schmidt G., Hommel U.;

RT "Structural basis for LFA-1 inhibition upon llovasstatin binding to the CD11a I-domain.";

RL J. Mol. Biol. 292:1-9 (1999).

CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2, ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES AND MONOCYTES.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L ASSOCIATES WITH BETA-2.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS: 2; Event=Alternative splicing; Named isoforms=2; Name=1; IsoId=P20701-1; Sequence=Displayed; Name=2; IsoId=P20701-2; Sequence=VSP 002738; Note=No experimental confirmation available;

CC -1- TISSUE SPECIFICITY: LEUKOCYTES.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: Contains 1 VWFA domain.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry; WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Y00796; CAA68747.1; -.

DR EMBL; AC002310; AAC31672.1; -.

DR PIR; S03308; S03308.

DR PDB; 1LFA; 29-JAN-96.

DR PDB; 1ZON; 07-DEC-96.

DR PDB; 1ZOO; 07-DEC-96.

DR PDB; 1ZOP; 07-DEC-96.

DR PDB; 1COP; 07-AUG-00.

DR PDB; 1DGO; 03-FEB-00.

DR PDB; 1MJN; 28-JAN-03.

DR PDB; 1MQ9; 14-JAN-03.

DR PDB; 1MQA; 14-JAN-03.

DR Genew; HGNC:6148; ITGAL.

DR MIM; 153370; -.

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0006928; P:cell motility; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS00234; VWFA; 1.

DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane; Signal; 3D-structure; Magnesium; Calcium; Repeat; Alternative splicing.

FT SIGNAL 1 25

FT CHAIN 26 1170 INTEGRIN ALPHA-L.

FT DOMAIN 26 1088 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1089 1112 POTENTIAL.

FT DOMAIN 1113 1170 CYTOPLASMIC (POTENTIAL).

FT REPEAT 92 91 FG-GAP 1.

FT REPEAT 42 149 FG-GAP 2.

FT DOMAIN 170 349 VWFA.

FT REPEAT 2 2

FT REPEAT 401 455 FG-GAP 3.

FT REPEAT 457 516 FG-GAP 4.

FT REPEAT 518 575 FG-GAP 5.

FT REPEAT 578 630 FG-GAP 6.

FT REPEAT 630 630 FG-GAP 7.

FT CA_BIND 468 476 POTENTIAL.

FT CA_BIND 530 538 POTENTIAL.

FT CA_BIND 590 598 POTENTIAL.

FT SITE 1115 1119 GFFKR MOTIF.

FT DISULFID 73 80 BY SIMILARITY.

FT DISULFID 111 129 BY SIMILARITY.

FT DISULFID 653 707 BY SIMILARITY.

FT DISULFID 771 777 BY SIMILARITY.

FT DISULFID 845 861 BY SIMILARITY.

FT DISULFID 998 1013 BY SIMILARITY.

FT DISULFID 1021 1052 BY SIMILARITY.

FT CARBOHYD 65 N-LINKED (GLCNAC...) (POTENTIAL).

OX NCBI TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91268576; PubMed=2051027;
 RA Kaufmann Y., Tseng E., Springer T.A.;
 RT "Cloning of the murine lymphocyte function-associated molecule-1
 alpha-subunit and its expression in COS cells.";
 RL J. Immunol. 147:369-374(1991).
 RN [2]
 RP SEQUENCE OF 24-42.
 RX MEDLINE=85188276; PubMed=3887182;
 RA Springer T.A., Teplow D.B., Dreyer W.J.;
 RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
 glycoproteins and unexpected relation to leukocyte interferon.";
 RL Nature 314:540-542(1985).
 CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
 ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
 INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
 MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
 AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
 SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
 LEUKOCYTES RECRUITMENT.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
 ASSOCIATES WITH BETA-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWF DOMAIN. INTEGRINS
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 1 VWF domain.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M60778; AAA39426.1;
 DR PIR; I56126; I56126.
 DR HSP; P20701; 1LFA.
 DR MGD; MGI:96606; Itgal.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VFWA; 1.
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 Signal; Magnesium; Calcium;
 KW Repeat.
 KW SIGNAL
 FT CHAIN 1 23
 FT DOMAIN 24 1163
 FT TRANSMEM 24 1084
 FT TRANSMEM 1085 1108
 FT DOMAIN 1109 1163
 FT REPEAT 39 88
 FT REPEAT ? ?
 FT DOMAIN 148 334
 FT REPEAT ? ?
 FT REPEAT 399 454
 FT REPEAT 455 514
 FT REPEAT 516 573
 FT REPEAT 576 628
 FT REPEAT 466 474
 FT CA_BIND 466 474
 FT CA_BIND 528 536

FT CA_BIND 588 596
 FT SITE 1111 1115
 FT DISULFID 70 77
 FT DISULFID 108 126
 FT DISULFID 147 199
 FT DISULFID 651 705
 FT DISULFID 767 773
 FT DISULFID 840 856
 FT DISULFID 993 1009
 FT DISULFID 1017 1048
 FT CARBOHYD 86 86
 FT CARBOHYD 185 185
 FT CARBOHYD 270 270
 FT CARBOHYD 444 444
 FT CARBOHYD 668 668
 FT CARBOHYD 696 696
 FT CARBOHYD 724 724
 FT CARBOHYD 728 728
 FT CARBOHYD 776 776
 FT CARBOHYD 857 857
 FT CARBOHYD 880 880
 FT CARBOHYD 890 890
 FT CARBOHYD 899 899
 FT CARBOHYD 927 927
 FT CARBOHYD 1056 1056
 SQ SEQUENCE 1163 AA; 128343 MW; A7A3078489E8232F CRC64;
 Query Match 26.28; Score 1538.5; DB 1; Length 1163;
 Best Local Similarity 34.28; Pred. No. 3.6e-96;
 Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;
 QY 1 FNLDTENAMTFQENA-RGFQSVVVLQSGSRVVGAPQEIIVAAQNRGLSYCDYSTGCEP 59
 DB 24 YNLDTRTQSLAQAGRHFGYQVLQIEDG-VVVGAPCE---GNDTGLYHCRTSSEFCQP 79
 QY 60 IRLQVPVEAVNMSGSLAATSPQLLACGPTVHQTCSNTYVYVKGCLFLFGSNLRQOPQ 119
 DB 80 VSLH-GSNHTSKVLYGMTLATDAAGSLACDPLGSLRTCDQNTYLSGLCYLFPQSLGPM 138
 QY 120 KFPALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEPISTVMEQLKKSKTLFSLMOYSEE 179
 DB 139 QNRPAYQECMKGRKVDLVFLFDGSGSLDRKDFEKLFEKMDVMKRLSNVTSYQFAAVQSTD 198
 QY 180 FRIHFTKEF-QNNPNPRSLIKPITQLLGRTHATGIRKVVRELNFNTNGARKNAFKILI 238
 DB 199 CRTEFTLDYVKNQKNDVLLGSGVQPMFLTNTRFVYVAVHFKESGARDPATKVLV 258
 QY 239 LIIDGKEGDPGLGYEDVIPLEADREG-----VIRYVIGVDAPRSKSKOELNTVASKP 291
 DB 259 IITDG-----EASDKGNISAAHDITRYIIGICKHFVSQKQKTLHIFASEP 304
 QY 292 PRDHVFOINNFEALKTIQNLREKI FAIEGTQTSSSSSSEHEMSQEGFSAITNSGPLLS 351
 DB 305 VEEFVKILDTFEKLUKDLFTDLQRIYAIEGTNRQDLTSFNWELSSSGISADLSKGHAVVG 364
 QY 352 TVGSDYDWAGGVF-LYTSKEKSTFTINMTRVDSMDMDAYLGAAA-IILNRNVQSLVLGAPR 409
 DB 365 AVGAKDWAGGFLDLREDLQGATFVQEPFLTSVGRGGLYGYTVAMWTSRSSRPLLAACAPR 424
 QY 410 YQHIGLVAMFR--QNTGMWESNANVKGTQIAGYFGALSCSVVDVDSNGSTDVLVIGAPHY 467
 DB 425 YQHVQVLLFQAPGAGGRWNQTKIEGTQISYFGGELCSVDLDQDGEAEALLIGAPLFF 484
 QY 468 EOTGGQVSVCPPLPRGORARWQCDVLYGEGQPMGRFGAALTVLGVNKGDKLTDVAIGA 527
 DB 485 GEQGRGVFTY---QRRQSUFEMVSELOQDGPYPLGRFGAALTALTDINGRLTLDVAIGA 541
 QY 528 PGEEDNRGAVYLFHGTSGSGISPSHSORISAGSKLSRLQYFGQSLSGQDLTMDGLVDLT 587
 DB 542 PLEE--QGVYVIFNGKPG-GLSPQSPQRIQAGVFPGRVWFGRIHGVKDLGGDRLDV 598
 QY 588 VQAQHVLLRSQVPLRVKALMEFNPREAVNPFECDNQVVKYKAG-EVRVCLVHVKST 646

FT CARBOHYD 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128983 MW; 88331C115DCCCFD CRC64;

Query Match 19.8%; Score 1161.5; DB 1; Length 1167;
Best Local Similarity 29.0%; Pred. No. 1.4e-70;
Matches 357; Conservative 214; Mismatches 457; Indels 205; Gaps 43;

QY 1 ENLDTENA--MTFQENARGFCQSVVQLOGSRVVVGAQOEIVAAANQSGS-----LYCODY 52
DB 20 FNMVDVMAWVATLQPGAPAVLSSLLHLDPNS-----NOTCLLVARRSSNRNTAALYRCAL 74
QY 53 STGSCERIRLQVPEAVANMSLGLSLAATT--SPQLLAC-GPTVHQTCSENTYVKGCLFL 109
DB 75 SI-SPDETACO-PVEHICMPKGRYQGVTLVGNHGVLCVQVQAKRSLNSELTGACSL 132
QY 110 FGSNLRQOPKPFPEALRG-----C-----PQE 131
DB 133 LTPNLDLQAQYFSDLEGFLDPGAHVDSGDYCRSGKSGTGBEKSARRRRVTEDEEED 192
QY 132 DSDIAFLVDGSGSIIPHDFRAKPSISTVMEQL--KSKTLFSLMOYSEEPRIHFTPEF 189
DB 193 GTEIAFLVDGSGSGPSDFQAKNPISMTMNFVEKCECFNPFALVQVGAIVTQEDFQES 252
QY 190 QNNPNRSLIKPITQLGRTHATGIRKVRLEFNIINGARKNAFKILILITDGEKFGDP 249
DB 253 RDINASLAKVOSIVOVKEVKTASAMQHVLDNIPIPSGRSKAKLVWVLTGDIGDP 312
QY 250 LGYEDVPEADREGVIRVIGVDGDAFRSEKSRQELNTVASKPRPDHVFQIINNFEALKTQ 309
DB 313 LNLTVINSKMGQVVRFAIGVDRFNKNTVRELKLIASDPKEAHTKVTNYSALDGLL 372
QY 310 NQLEKRIEIAETQTGSSSEFHEMSQSFNAITSNGP-LLSTVGSVDWAGGVFLY-TS 367
DB 373 SKLOQRIVHMEGT--VGDALQYLAQTGFSAQILDKGOVLLGTVGAFNWSGGALLYSTQ 429
QY 368 KEKSTFIINMT-RVDS-DNMDAYLGYAAAIILNRVQSLVGLAPRYQHILGLVAMFRONTGM 425
DB 430 NCRGFLNQTKADESRITQVSYLGLSVLAVLHKAGISVAGAPRHLKAGVFELEKDR- 488
QY 426 WESNA---NVKQTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSVCPILR 482
DB 489 -EEDAFVRRIEGQMSYFGSVLCPVDIDMDGTTDFLLVAAPFYHIRGEGRVYVQYPE 547
QY 483 GQARWOCDAVLGYGQGPWGFGAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSGHPGLTNSRFGPAMAAGVINDQKFTDVAIGAPLFGFGAGDGASYS 606
QY 537 VTLFHTGSGISPSHSORISAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLL 596
DB 607 VYIYNGHSG-GLYDPSQOIRASSVASGLHYFGMSVSGGLDPNGDGLADITVGSRSADV 665
QY 597 LRSQVLRVKAIMEPNREVARNVFECDNQVVKGEAGEVRVCLHVQKS---TRDRLEG 653
DB 666 LRSRPVOLDTVSMTEPT-----DALPMVFIRGM--DVNLCFEVDSSVVAEPGLREM 715
QY 654 QIQSVVYTDLALDSRPHSVAFNETKNSRTOQVILGLTQFC----- 696
DB 716 FLNFTVDV-----TKQRQLQCEDSGSCQCLRKWNGSGFLCBHFWLI 760
QY 697 ETCLKLQPCIEDPVPVILNLSLVTPLSAFNGLR-----PVLAEADQRLLFTALF--P 750
DB 761 STEEL-----CEDCFSNITIKTYE-----FOTSGRRDYPNPTL--DHYKEPSAIFQLP 809
QY 751 FEKNGNDNICODDLSITFSFMSLCLVVGGRPRENVTVTVNNGEDSVRTQVTFPPFLD 810
DB 810 YEKQCKGNKVFCAEIQLTATN-ISQBELVVGVTKEVTMNLISLNSGDSYMTNMAALNPRN 868
QY 811 LSVRKVSTLQONQRSRPLACESASSSTVSALKSTSCSINHPIFPENSEVTFNITFDV 870
DB 869 LQFKKI-----QKPSRPVQDCDDPKPV---ASVLVNMCKIGHPIIL-KRSSVNVSVTQWL 918
QY 871 DSKASLGNKLLKANVTSENNPRTNKTEFQLELPVKYAVYVWVTVSHGVSTKYLNFPTASE 930

DB 919 EESVFNRTADITVTISNSNEKSLARETR---SLQPRHAFIAVLSR--PSVMTMN--TSQ 971
QY 931 NTSRVNMQHOYVSNLQORSIPISLVFLVPLVNLQNTVIMDRPOVTPSPENLSST-----CHT 985
DB 972 SPDSHKEFFNVHGENLFGAVFQLOICVPIKQDP-----QIVRVNLTKTQDHTECTQ 1025
QY 986 KERLPSHSDFLAEIRKAPVNVCSIAVCORIQCDIPFPGIQEEFNATLKNLSFDWYIKTS 1045
DB 1026 SQSPACGSDPVQVHKMSHVCAI-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLLIVSTA-----EILFNDVSFTLLPQCGAFVRSQTETKVEPF-----EVPNPLPLIV 1095
DB 1066 HTRQLLRDVSLEPILGEISFNKSLYBGLNAE-----NHRKITVIFLKEEETRSPLII 1119
QY 1096 GSSVGGLLALALTAALYKLGFFKROYKDMWSE 1128
DB 1120 GSSIGGLLVLLVIAILFKCGFFRKYQQLNLE 1152

RESULT 8
ITAE HUMAN
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and Lymphocytes;
RX MEDLINE=94164962; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B.,
Parker C.M.;
RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Aniketer Y., Dietrich N.L., Maduro V.V., McDowell G.,
Shoclerauk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,
Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
RN [4]
RP MUTAGENESIS OF ASP-109 AND PHE-316.
RX MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-B/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN
CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL
CC EPITHELIAL CELLS.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L25851; AAB59359.2; -
CC EMBL; AF168787; AAF43107.1; -
CC PIR; A53213; A53213.
CC HSSP; P11215; LA8X.
CC Genew; HGNC:6147; ITGAE.
CC MIM; 604682; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 3.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Repeat; Polymorphism; Magnesium; Calcium.
CC SIGNAL; 1 18
FT CHAIN 19 1179 INTEGRIN ALPHA-E.
FT CHAIN 19 177 INTEGRIN ALPHA-E LIGHT CHAIN.
FT CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
FT CHAIN 19 1124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1125 1147 POTENTIAL.
FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 181 198 GLU-RICH (ACIDIC).
FT REPEAT ? ? FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT REPEAT 145 199 X-DOMAIN (EXTRA DOMAIN).
FT DOMAIN 200 391 VWFA.
FT REPEAT 401 456 FG-GAP 3.
FT REPEAT 457 506 FG-GAP 4.
FT REPEAT 510 571 FG-GAP 5.
FT REPEAT 573 638 FG-GAP 6.
FT REPEAT 641 693 FG-GAP 7.
FT CA_BIND 522 530 POTENTIAL.
FT CA_BIND 586 594 POTENTIAL.
FT SITE 654 662 GPFKR MOTIF.
FT SITE 1150 1154 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 126 159 BY SIMILARITY.
FT DISULFID 706 762 BY SIMILARITY.
FT DISULFID 823 829 BY SIMILARITY.
FT DISULFID 893 907 BY SIMILARITY.
FT DISULFID 1008 1033 BY SIMILARITY.
FT DISULFID 1041 1057 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	782	782	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	934	934	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	954	954	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1065	1065	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1096	1096	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	360	360	D -> E.
FT	VARIANT	1041	1041	/FTID=VAR_008884.
FT	MUTAGEN	208	208	/FTID=VAR_008885.
FT	MUTAGEN	316	316	D->A: LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	477	477	E->A: LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	482	482	V -> I (IN REF. 3).
FT	CONFLICT	482	482	Q -> R (IN REF. 3).
FT	CONFLICT	950	950	R -> W (IN REF. 3).
FT	CONFLICT	1019	1019	A -> V (IN REF. 3).
SQ	SEQUENCE	1179 AA; 130088 MW; E558902EDP9D95E1 CRC64;		

Query Match 19.6%; Score 1153; DB 1; Length 1179;
Best Local Similarity 29.1%; Pred. No. 5.3e-70;
Matches 342; Conservative 214; Mismatches 449; Indels 172; Gaps 39;

QY	45	GSLYOCYSTGS--CEPI-RLOVP-----VEAVNMSGLSLAATTSPPOLLACCGPTVHQ	95
DB	65	GPLHRCSLVQDEILCHPVEHVPPIKGRHGRVTVVRSHHGVLICI-----QVLVRRP--HS	117
QY	96	TCSENTYVKGCLFGLSGLNRQPOQ-----SDIAFLVDGSGSIIPHDPRRAKEFISTVMEQL--	119
DB	118	LSSELT--GTCSLGPDRLRPOAANFFDENLDPDARVDGDCYSNKEGGEDDVNTA	174
QY	120	KPEALRGCPQED-----SIAFLVDGSGSIIPHDPRRAKEFISTVMEQL--	164
DB	175	RORALEKEEEDDEEEDDEEAGTEITAILDGGSGIDPPDFORAKDFISNMNRNFEY	234
QY	165	KSKTFLSLMOYSEEFRIHFTKFEONNPNRSLIKPIQLLGRTHATGIRKVVRELEN	224
DB	235	KCFECNFALVGVGVIGTQTEFLRDSQDVNASLARVQNTQVGSVTKTASAMOHVLDISFT	294
QY	225	ITNGARKNAFKILILITDGEKFGDPLGVEDVPEADREGVIRYVIGCVDAFSEKSRQEL	284
DB	295	SSHGSRKASKVMVVLTDGGIFEDPLNLTTVINSPKMQGVVERFAICVGEFSAKARTREL	354
QY	285	NTVASKPRDRHVQINFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGSAAIT	344
DB	355	NLIASDPDETHAFKVTNYMALDGLLSKLRYNIISMEGT---VGDALHYQLAQIGFSAQIL	411
QY	345	SNGP--LLSTVGSYDWDAGGVFLY--TSKEXSTFINMTFVDSMDMA----YLGVAAILLRN	398
DB	412	DERQVLLGAVGAFDWSGGALLYDTRRRGRFLNQTAATAAADAAEAQSYLGVYAVAVLHKT	471
QY	399	RVQSLVLGAPRYQHI GLVAMFR--QNTGMWESNANY--KGTQICAYFGASLCSVDVDSNGST	456
DB	472	CSLSYVAGAPQYKHG--AVFELQKEGREASFLPVLGEQMGSYFGSELCPVIDDMDGST	529
QY	457	DLVLIGAPHYBQTRGGQVSCPLPRGQRARQCDAVLVGEQGPWGFGAALTVLGDVN	516
DB	530	DFLLVAAPFYHVHGEGRVYVYRLSE--QDGSFSLARILSHRGFTNARFGFAMAAGDLS	588
QY	517	GDKLTDVAIGAP----GEEDNR--CAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQ	570
DB	589	QDKLTDVAIGAPLEGFGADDDGASFGSVIYNG--HWDGLSASPSQIRASTVAPGLQYFGM	647
QY	571	SLSGGQDLTMDGLVDLTVGAQGHVLLLRLOPVLKAIWFEFNPVARNVFECDQVVKG	630
DB	648	SNAGFDISGDLADITVGTLCQAVFVRSPVVRKLVKSNAPFTSALP-----IGF	697
QY	631	KEAGEVRVCLHVOKSTRDLREGQIQSVVYTDALDGRPHGRVAFNETKNSRTRQTVL	690
DB	698	NGVNVRLCFEI--SSVTTASESGLEALLNFITLDVDVGKRRRLQCSRVSRCLGUREWS	756
QY	691	GLTQTCETKLQLPN-----CIEDPVPVLRNFSLVGTPLSAFGLNLRPLVLAEDAQRFT	746
DB	757	SGSQLCEDL--LLMPTEGELCEDCFSNASKVYSYL--QTPEGQTOHPQILDRYTPFAI	814

Qy 747 ALFFPKCKGNDNTCDDLSITFSFMSLDCLVCGPREFNVTVTVRNDGEDSYRTQVTFP 806
Db 815 FQLPYEKACKNKJFCVAELQLA-TTVSQQLWGLTKELTNLNLNSGDSYMTSMALN 873
Qy 807 PPLDLISYKVTQNSORSORWRACASASTEVSGALKSTSCSINHPPIFENSEVTFNI 866
Db 874 YPRNLQ-----LRMKQPPSPNIQCDDPQV---ASVLNMCNIGHPVL-KRSASHSV 923
Qy 867 TFDVDSKASLGNKLLKANVTSENN---MPRTNKTBFQ---LELPVKYAVYVTVTSHGV 919
Db 924 VWOLEENAFPNRTADITVTVNSNERRSLANETHTLQPHGFVAVLKSPSMYVNTGQL 983
Qy 920 S--TKYLNFTASENTSRVQHQYQVSNLQGRSLPISLVPLVPLVPLVQVWDRQVTFSE 977
Db 994 SHHKEFLPHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAAVKLLTRTQ 1028
Qy 978 NLSSTCTKHPKSHSDPLAELRKAPVNVVNCIAVCQRIQCIPFFGQEEFNATLKGMLS 1037
Db 1029 ASTVCTWSQERACAYSS-VQVHEWHSVCVIA-----SDKENVTVAEIS 1073
Qy 1038 FDMVYKTSNHLIVST-----AEILFNDVSFTLLPQOGAFVRSQTKETKVEPEVNPL 1091
Db 1074 WD-----HSELLKDVTELQILGEISFNKSLYEGLNAENH--RTKITVWELKDEYHSL 1125
Qy 1092 PLIVGVSSVGGILLALLIATAALYKLGFFKRYQKDMWSE 1128
Db 1126 PIITKGVSGLLVLIVLILVILFKCGFFKRYQQLNLE 1162

RESULT 9

ITAL HUMAN
ID ITAL HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit";
RL J. Biol. Chem. 268:2989-2996 (1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm".
CC PIR; A45226; A45226.
DR PDB; 1OC5; 17-MAY-00.
DR Genew; HGNC:6134; ITGAI.
DR MIM; 192968;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; NAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.

DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; wva; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; WVA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; WVA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure.
KW DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1114 1136 POTENTIAL.
FT DOMAIN 1137 1151 FG-GAP 1.
FT REPEAT 16 75 FG-GAP 2.
FT REPEAT 16 75 FG-GAP 2.
FT DOMAIN 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA BIND 470 478 POTENTIAL.
FT CA BIND 552 560 POTENTIAL.
FT CA BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GPFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;

Query Match

Best Local Similarity 18.8%; Score 1103.5; DB 1; Length 1151;
Matches 344; Conservative 212; Mismatches 486; Indels 195; Gaps 44;

Qy 1 FNLDTENAMTFQENARG-FGQSVVQL----QGSVVVVGAPQBIIVANORGSLYQCDYSTGS 56
Db 1 FNVDVKNSTSGPVEDMFGYTVQYENECKWVLIGSLPVGQPKNRTGVDVYKCPVGRGE 60
Qy 57 CEP-IRLQVPEA-----VNMSLGLSLAATSPQLLACGTVTQTCSENTYVVKGL 106
Db 61 SLPCVKLDLPVNTSIPNVTENKNTFGSTL-VTNPNGGFLACGPLYAIRCGLHYTTGI 119

QY 107 CFLFGSNLRQOPKFPALRGCPQEDSDIAFLVDGSGSIIPHDPRAKBFISIVMBQLK- 165
DB 120 CSDVSPFTQVNSIAP--VOECSTQ-LDIVILVDGNSIYPPWDSVTA--FLNDLLKRMID 174
QY 166 -KSKTLFSLMOYSEEFRIHFTKFEFONNPNRSLIKIPITQLLGR-THTATGIRKVVRELF 223
DB 175 GPKQTQVGIQVGENVTHEFNLNKYSSTEVLVAAKIVQGRQTMALTGDTTARKEAF 234
QY 224 NITNGARKNAFKILITDGEKFGDPLGYEDVIPEADREGVIYVIGVGDAFR-----SE 278
DB 235 TEARGARRGVKKVMVITDGEH-DNHLKVKVQDCEDENIQRFSAIILGYSYNRGNLSTE 293
QY 279 KSRQELNTVASKPRPHVQIINNFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQBG 338
DB 294 KFVEEIKIASEPTEHFFNVSDALVITVKTGERIFALEATADQSAASEMEMSQTG 353
QY 339 FSAAITNSGPLLSTGVSXWAGGVFLYTSKE-----KSTP-INMTVDSMDNDVLYGAA 392
DB 354 FSAHYSQDWMVLGAVGAYDNGVTVMQASQIIIPRNTTFNVESTKKNPL-ASYLGVTV 412
QY 393 AITLNRVOSL-VLGAAPRYQHIGLVAMFRONTGMESNANVKGTQIGAYFGASLCSVDVD 451
DB 413 NSATASSGDVLYTAGQPRYNTHTQVVIYRMEDGNKILQTLSEQIGSYFGSILITTTDID 472
QY 452 SNGSTDLVLGAPHY-----YEQTR-GGQVSVCPPLPRQARWQCDVAVLY 495
DB 473 KDSNTDILLVGPAPMYNGTEKEQGVVYVALNQTREYQMSLEPIKQTCSSRQNSCTT 532
QY 496 GEOGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLPHGTSGSGISPSHSO 554
DB 533 ENKNEPCGARFGTAIAVADNLDGDFNDIVIGAPLEDHGGVAYIYHG-SGKTIKRYAQ 591
QY 555 RIAGSKLSRLQVFGOSLGGQDLTMDGLVDLTGAGQGHVLLRSOPVLVRKAIMFNPR 614
DB 592 RIFSGGDKTLKFFGSGIHGEMDLNGDGLTDTVITGLGGALEFWSRDAVAVKVTWVFN 651
QY 615 EVARNVFECDQVVKQKAG--EVRVCLHVQ-KSTRDLRREGIOQSVVTDIALDSGRPH 671
DB 652 KVNIOKKNCH--MEGKETVCINATVCFEVLKSKEDTIYEADLQ----YRVLDSLRQI 704
QY 672 SRAVENET-----KNSTRQVQLGTOTCETLKLQPLNCIEDPSPVIVLBNLSVCT 725
DB 705 SRSFSGTQERKQVRNITVKSEC-----TKHSFYMLDKHDFQDSYR----ITLDFNLT-D 755
QY 726 PLSAFNLRPVLAEDAQRILFTALFPPEKNCNDNICDDLSITFSFMSLDCLVGVGPRE- 784
DB 756 PENG-----PVLDDSLPNSVHEIYIPAKDGNKEKICISLSLHVATTEKDLIVRSQNDK 810
QY 785 FNVTVTRNDGDSYRTQTFPPLDLSYRKVSTLQNRQSRQSWRLACASASTEVSGAL 844
DB 811 FNVSLTVKNTKDSAYNTRIVHYSPLNLFSGTEAIQKD-----SCSN----- 853
QY 845 KSTSCSINHPIPENSEVTFNITFDVDSKASLGN-KLLKANVTSENMPRNTKTEFQLE 903
DB 854 HNTIKGVGFPLRGBMVTFKLQFNTSYLMENVITLYLSATSDSEPPETLSDNVNITS 913
QY 904 LPVKYAVYVMTVSHGVSYTKLNTFASENTSRVMOHQYVSN-----LGORS-----L 950
DB 914 IPVKVEVGLQFYS-SASEHISIAANETVPEVINSTEDIGNEINIPYLIRKSGSPMPLEL 972
QY 951 PISLVF-----LVVRLNQTQVIMDRPQVTFSENLSSTCHTK-----RLPS 991
DB 973 KLSISFPNNTSNGYPVLYFTGLSS-----SENANCRPHIEDFDSINSKGKMTT 1021
QY 992 HSDFLAELRKAPVNCVSIACORIQDIPFFGIQE-----EFNATUK 1033
DB 1022 STD-----HLKRGITLDCNTCKFATITCNLTSSDISQVNVSLILWKPTFKYSFSSNLTR 1078
QY 1034 GNLSPDWYIKTSHNHLIIVSTAILPNDVSFTLLPGGAFVRSOTETKVEPEVFNPLPL 1093
DB 1079 GEL-----RENASLVLSNN-----QKRELAIQISKDGLPGRVPL 1114
QY 1094 --IVGSSVGGLLLLALITALYAKLGFPRQYKDMSE 1128

DB 1115 WVLLSFAFAGLLMLLILALWKIGFFRPLKMKMEK 1151
RESULT 10
ITAH HUMAN
ID ITAH HUMAN STANDARD; PRT; 1189 AA.
AC Q9UKX5; Q9UKQ1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-11 precursor.
GN ITGAI1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=99417678; PubMed=10486209;
RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
Wang S.-X., Morris C.M., Krissansen G.W.;
RT "Cloning, sequence analysis, and chromosomal localization of the novel
human integrin alpha11 subunit (ITGAI1).";
RL Genomics 60:179-187(1999).
RN [2]
SEQUENCE FROM N.A.
RC TISSUE=Fetal muscle, and Uterus;
RX MEDLINE=99395147; PubMed=10464311;
RA Velling T., Kutsche-Gullberg M., Sejersen T., Gullberg D.;
RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)
Integrin. A collagen-binding, i domain-containing, beta(1)-associated
integrin alpha-chain present in muscle tissues.";
RL J. Biol. Chem. 274:25735-25742(1999).
RN [3]
SEQUENCE OF 954-1188 FROM N.A.
RC TISSUE=Fibroblast;
RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
CC FETAL MUSCLE CELLS (IN VITRO).
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF109681; AAF01258.1; -
DR EMBL; AF137378; AAD51919.2; -
DR EMBL; AL359064; CAB94392.1; -
DR HSSP; P17301; IAOX.
DR Genew; HGNC:6136; ITGAI1.
DR MIM; 604789; -
DR GO; GO:0008305; C:integrin complex; TAS.

DR GO: GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO: GO:000518; F:collagen binding activity; TAS.
DR GO: GO:0007160; P:cell-matrix adhesion; TAS.
DR GO: GO:0007517; P:muscle development; TAS.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR01185; INTEGRIN.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; FALSE_NEG.
DR PROSITE: PS0234; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium; Polymorphism.
FT SIGNAL 1 22
FT CHAIN 23 1189
FT DOMAIN 23 1142
FT TRANSMEM 1143 1165
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT FG-GAP 1. 94
FT FG-GAP 2. 163
FT FG-GAP 3. 163
FT FG-GAP 4. 163
FT FG-GAP 5. 163
FT FG-GAP 6. 163
FT FG-GAP 7. 163
FT POLY-LEU. 1174 1177
FT CA_BIND 488 496
FT CA_BIND 551 559
FT CA_BIND 613 621
FT DISULFID 76 83
FT DISULFID 121 139
FT DISULFID 129 159
FT DISULFID 659 668
FT DISULFID 674 729
FT DISULFID 781 787
FT DISULFID 881 893
FT CARBOHYD 82 82
FT CARBOHYD 95 95
FT CARBOHYD 291 291
FT CARBOHYD 331 331
FT CARBOHYD 358 358
FT CARBOHYD 449 449
FT CARBOHYD 462 462
FT CARBOHYD 528 528
FT CARBOHYD 642 642
FT CARBOHYD 694 694
FT CARBOHYD 857 857
FT CARBOHYD 894 894
FT CARBOHYD 973 973
FT CARBOHYD 1032 1032
FT CARBOHYD 1040 1040
FT VARIANT 433 433
FT VARIANT 524 524
FT VARIANT 972 972
FT VARIANT 1003 1003
FT VARIANT 1030 1030
FT VARIANT 1094 1094
FT SEQUENCE 1189 AA; 133609 MW; 6030308A44CD52 CRC64;
/FTid=VAR_009889.
R -> L.
/FTid=VAR_009890.
L -> P.
/FTid=VAR_009891.
I -> M.
/FTid=VAR_009892.
MISSING.
/FTid=VAR_009893.
L -> V.
/FTid=VAR_009894.
MW; 6030308A44CD52 CRC64;

Query Match 18.6%; Score 1093.5; DB 1; Length 1189;
Best Local Similarity 28.2%; Pred. No. 5.9e-66;
Matches 348; Conservative 215; Mismatches 502; Indels 167; Gaps 47;

QY 1 ENLDTENAMTFOENARG-FGQSVVO--LQGSR-VVVGAPQEIIVAAQNRSLXQCDYSTGS 56
DB 23 FMDTRKPRVIFGSRITAFYVYQHQHDSGNKWLWVGAPLETNGYOKTGDVTKCPVIHGN 82
QY 57 CEPRL-----QVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVYKGLCFPL 110
DB 83 CTKLNLRVTLGNVSEKDNMRGLSLATNPKNDSFLACSPLWSHEGSSYVTTGCSR 142
QY 111 GSNLRQOPKPFPEALRGCPQEDSDTAFLVDGSGSIIPHDFRRAKEFISTVMEQ--LKKSK 168
DB 143 NSNFRFSKTAP-ALQRC-QTYMDIVIVLDGNSIYP--WVEVQHFLNLIKFKYIGPOQ 198
QY 169 TLPFLMYSSEFRIHFTKFEFQNNPNRSLIKPITOLLG-RTHATGIRKVVRELFINITN 227
DB 199 IQGVVVQYGEDVHFEHLNDYRSKDVVEAAASHIEQGGTETRTAFGIFARSEARQ--K 256
QY 228 GARKNAFKILILITDGEKFGDPLGYEDVIPEADREGVIRYIGV-----GDAPRSKSRQ 282
DB 257 GGRKGAKKVMIVITDGHSDSP-DLEKVIQSERDNVTRVAVAVLGYVYNNRGINPETFLN 315
QY 283 ELNTVASKPDRHVFOINNPEALKTIONQIRKIFAIEGTQTGSSSSPHEHMSQEFSA 342
DB 316 EIKYIASDPDDKHFFNVTDAAALKOIVDALGRIPSLGNTK-NETSPGLEMSQTGFSSH 374
QY 343 ITSNGPLLTSGSYDMAGGVFLVTSKEK-----STEINMTRVDSMDNDAYLVGAAAII 397
DB 375 VVEDGVLLGAVGAYDNWGAVALKETSAGKVIPLRESYLKEPPELKHGAYLGVTVTSVVS 434
QY 398 NRY-QSLVLGAPRYQHIGLVAMF-RQNTGMESNANVKTQIGAYFGASLCSDVDVDSNGS 455
DB 435 SRQGRVYVAGAPRFNHTGKIVLFTMNRSLTIHQAMRGQIGSYFGSEITSDIDGDGV 494
QY 456 TDVLVLGAPHYEQTR-GQVSVCPPLGRQARWQCDVLYGEQGPWGRFGAALTVLGD 514
DB 495 TDVLLVGAAPYFNEGRGKVVYEL---RQNRVYNGTLKDSHSYQNARFGSSIASVRD 551
QY 515 VNGDKLTDVAIGAPGEDNKGAVLPHGTSGSGISPSHSORIASGKSLSPRLQYFGOSLSG 574
DB 552 LNQDSYNDVVVGAPLEDNHAGAIYIFHGFRGS-ILKTPKQRTASELATGLQYFGCSIHG 610
QY 575 GQDLTMDGLVDLTVGAQGHVLLRSQVLRVKAIMEFNPREVARNVF--ECNDQVVKGKE 632
DB 611 QLDLNEGLIDLAVGALGNVILWSRPVQINASHLHPEPSKI-NIFHRDC-----XR 661
QY 633 AGEVRVCL-----HVQKSTRDLREGQIQSVVYVYDLADSDGRPHRAVNET 679
DB 662 SGRDATCLAAFLCFTPIFLAPHFQTTVG-----IRYNATMDERRYTTPRAHLDEG 711
QY 680 KNS-TRRQTOVLGLTOTCETLKLQLPNCIEDPVSPIVLRNLNLSVGTPLSAFGLNRPVLA 738
DB 712 GDRFTNRAVLSSGOELCERINPHVLD-TADYVYKPTVFSVEYSLEDP-----DHGPM 764
QY 739 EDARLFTALFPPEKNCGNNDICQDDL-----STTFSP 771
DB 765 DGMPTTLRVSVPEFWNGNEDEHCVPLDVLDAESDLPAMBYCQVLRKPKAQDCSAVTLSP 824
QY 772 MSLDCLVVGGRPFNVTVTRNDGESYRTQVTFPPLDLSYRKVSTLQNRQSRSLA 831
DB 825 DTTVFIIESTRQRAVEATELENGENAYSTVLNISQSANLQF--ASLIQKEDSDGS--IE 880
QY 832 CESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKAMVTSENN 891
DB 881 CVNEER-----RLQKQVCNVSYPPFRAKAKAVRDLDFEP-SKIFLHLHLEIAGSDSN 934
QY 892 MPRTNKTE--FQLELPVKYAVVMVTSHGVSSTKY---LNFTAS--ENTSRVMQHYQVSN 944
DB 935 ERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYEVKLNSSLERYDGIQPPFCIFRIQ 994
QY 945 LQO---RSLPISLFLVPLVPLNQTVDWRPQVTFSENLSTC-----HTKERLSHSDFLA 997
DB 995 LGLFPIHIGIMMKITIPATRSNGRLKLRDLFT-DEVANTSCNIGWNSTEYRTPVE--E 1051


```
Db 370 VGFAEYSPQNNILMLGAVGAYDWSGTVVQKTPHGLIFSKQAFQEQILQDRNHSSVLYGS 429
Qy 392 AAIIILNRVQSLVGLAPRVQHGLVAMFRQNTGMESNANV-----KGTQIGAVFGASL 445
Db 430 VASISTGNSVHFVAGAPRANVTGQIVLYSVN-----ENGNTVVIOSQGDQIGSYFGSVL 484
Qy 446 CSVDVDSNGSTDLVLIGAPHYEQTR--GGQVSVCLPRGORARWQCDVAVLGEQGPWG 503
Db 485 CAVDVNDKDTIDVLLVGAQPMYNDLKKEGRVYLTITKG-TLNWH--QFLSGPGLENA 541
Qy 504 RFGAALTIVLDVNGDKLTDAVIGAPGEEDNRGAVLYFHTGTSGSGISPSHSQRIAGS--KL 561
Db 542 RFGSAIALSDINMDGFNDVIGVSPLENQNSGAVIYNGHEG-IRLRYSQKILGSDRAF 600
Qy 562 SPLQVFGSLSCGQDLTWGLVDLTGAGCHVLLRSQPVLRKAIMFNPREVARNVF 621
Db 601 SSHLQVFGRLDGYGDLNGSDITDVSAGAFQVQVQLWSQSIADVSVDASFTPKKI--TLL 658
Qy 622 ECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVTVYDIALD----SGRPHSRAVEN 677
Db 659 NKVAEI-----KLKLP-----SAKFRPNQNNQVAIVYNTITIDEDQFSRVSRLGPK 707
Qy 678 ETKNSTRRTQVGLGLTQCE--TLKLQLPNCIEDPVPSPVLRNLNPSL--VGTPLSAFGLN 733
Db 708 ENNERCLQKTMIVSQARCSSEYIIHQEPS---DIISPLNLCNLSLENPGT-----756
Qy 734 RPLVADDAQLFTALPFPKKNQNDNICDDLSITP-----SPMSLCLVGVGPREFNVTV 789
Db 757 NPALAEYSETKVFSIPFHKDCGDDGVCISDLVNLVQQLPATQQQPPVINSQNKRLTFSV 816
Qy 790 TVRNDGEDSVRTQVTFEPLDLSYRVKSTLQNRQSRWSRLACESASST--EVSGLKSTS 848
Db 817 QLNKKESAYNTIIVDFSENLF-----ASMSMPVDGTEVTCQIASSQKSVT 864
Qy 849 CSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMMPTNKTEFQLELPVKY 908
Db 865 CNVGYPAKSKQVTFITNFDFNLQ-NLQNAQASISPRALSQSEENMADNSVNLKSLLY 923
Qy 909 AVTVMTVSHGVSTKYLNFTASENTSRVMQHYQVSNLQR-----SLPISLVFLV 958
Db 924 DABIHIT-RSTNINFEVSLDGNVSSV-HSFE--DIGPKFIFSIVKTTGSVPVMSA---976
Qy 959 PVRLNQTIVDRPQVTFSEN--LSSTCTHKE-----RLPSSHDFLAE- 998
Db 977 -----SVIIHIIPOYTKDNPLMYLTVGVHTDQAGDISCEABINPLKIGTSSSVFSKEN 1030
Qy 999 LRKAPVNCIAVCORIQCDIPFGIQEENFATLKGNSLFDWYIKTSHNHLIVSTAETI- 1057
Db 1031 FRHikelNCRTASCSNIMCWLRLDQVKGVEYFLNVSTRIWNGTFAASTFTQVQLTAAAEID 1090
Qy 1058 LFNDSVFTL-----LPGQAFVRSQTETKVEPFE-VPNPLPLIVGSSVGGILLALITA 1110
Db 1091 TYNPQIVVIBENTVTP-----LTIKPKHEKVEVPTGVIVGSGVIAGILLALVA 1140
Qy 1111 ALVKGIGFFKRYQKDM 1125
Db 1141 ILWLKIGFFKRYKXDM 1155

RESULT 12
ITA2 MOUSE
ID ITA2_MOUSE STANDARD; PRT: 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID=10090;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha-
RT 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOLOGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VMPA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z29987; CAA82877.1; -
CC FIR; S44142; S44142.
CC HSP; P17301; LAOX.
CC MGD; MGI:96600; Icg2.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 4.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS00234; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 26
FT CHAIN 27 1178
FT DOMAIN 27 1129
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1130 1151
FT POTENTIAL.
FT DOMAIN 1152 1178
FT CITOPLASMIC (POTENTIAL).
FT REPEAT 42 100
FT FG-GAP 1.
FT REPEAT 7 7
FT FG-GAP 2.
FT DOMAIN 185 375
FT VMPA.
FT REPEAT 7 7
FT FG-GAP 3.
FT REPEAT 431 483
FT FG-GAP 4.
FT REPEAT 485 546
FT FG-GAP 5.
FT REPEAT 548 607
FT FG-GAP 6.
FT REPEAT 612 664
FT FG-GAP 7.
FT CA_BIND 496 504
FT POTENTIAL.
```

QY	680	KNSRTRQTVLGLTQTCT--LKLQLPNCIEDVSPIVLRNLPSLVGCTPLSAGFNLRPVL	737
DB	718	SRFLQKNVWVNEVQRCSHHISIQKPS--DVVNPLDLRVDISLENPGTS-----PAL	768
QY	738	AEDAQRFLTALPFPEKNCNDNLCQDLSI-----TFSPMSLDCLVVGPREFNV	788
DB	769	EAYSETVKVFSIPFYKEGSDGICISDLIDVQOLPAIQTSF-----IVSNQNKRLTFS	823
QY	789	VTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASST-EVSGALKST	847
DB	824	VILKNRGESAYNTVLAEFSENLF-----ASFSPVDGTEVTCVSGSQSV	871
QY	848	SCSINHPIIPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTE--FOLELP	905
DB	872	TCDVGYPALKSEQQVTFITNDFNLQ-NLQNOAAINQAFSESQ--ETNKADNSVSLTIP	928
QY	906	VKYAVVMVYTSCHGVSTKYLNLFASENTSRVMQHOYQVSNLQOR-----SLPISLV	955
DB	929	LLYDAELHLT-RSTNFINFEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSM	984
QY	956	FLV-----PVRNLQTVIWRDPQVTF-SENLS	980
DB	985	LVTIHTIPOYTKENKPLLTLTGTDQAGDISCTAEINPLKLPHTA---PSVSFKENFR	1040
QY	981	STCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIPFGIQIEFNATLKGNI	1040
DB	1041	---HTKE-----LDCRTTSCNITCWLKDLHMAEYFINVTVRVNRT	1080
QY	1041	YIKTSNHLILVSTAEILFNDSVFTLLPGCGAFVRSQTETKVPFVDPNPLIVGSSVG	1100
DB	1081	FAASTFQTVQLTAAAEIDHNPOLFVIEENAVTIPLIMKPKTEKAEVPT--GVIIGSII	1138
QY	1101	GLLLALITAAALKYLGFFKRYQKDM	1125
DB	1139	GILLLLAMTAGLWKLGFKKRYQKDM	1163

RESULT 13

ITAZ_HUMAN

ID	ITAZ_HUMAN	STANDARD;	PRT;	1181 AA.
AC	P17301;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)			
DE	(Collagen receptor) (VLA-2 alpha chain) (CD49b).			
GN	ITGA2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI	TaxID=9606;			
RN	[1]			
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.			
RC	TISSUE=Endothelial cells;			
RX	MEDLINE=89308879; PubMed=2545729;			
RA	Takada Y., Hemler M.E.;			
RT	"The primary structure of the VLA-2/collagen receptor alpha 2 subunit			
RT	(platelet GP1a): homology to other integrins and the presence of a			
RT	possible collagen-binding domain.";			
RN	J. Cell Biol. 109:397-407(1989).			
RN	[2]			
RN	SEQUENCE FROM N.A.			
RA	Rieder M.J., Amel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,			
RA	Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;			
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RN	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.			
RX	MEDLINE=98019223; PubMed=9353312;			
RA	Emley J., King S.L., Bergelson J.M., Liddington R.C.;			
RT	"Crystal structure of the I domain from integrin alpha2beta1.";			
RT	J. Biol. Chem. 272:28512-28517(1997).			
RN	[4]			
RN	VARIANT HPA-5 (BR)			

RX MEDLINE=94043762; PubMed=7901236;
RA Santoso S., Kalb R., Waika M., Kiefel V., Mueller-Eckhardt C.,
RA Newman P.J.;
RT "The human platelet allomorphs Br(a) and Brb are associated with a
RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
RT alpha 2).";
RL J. Clin. Invest. 92:2427-2432 (1993).
RN [5]
RP VARIANT GLU-534.
RX MEDLINE=20206009; PubMed=10741412;
RA Kroll H., Gardemann A., Fichter A., Haberbosch W., Santoso S.;
RA "The impact of the glycoprotein Ia collagen receptor subunit A1648G
RT gene polymorphism on coronary artery disease and acute myocardial
RT infarction.";
RL Thromb. Haemost. 83:392-396 (2000).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
CC associates with beta-1. Interacts with HPS5.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- POLYMORPHISM: POSITION 534 IS ASSOCIATED WITH PLATELET-SPECIFIC
CC ALLOANTIGEN HPA-5 (BR). HPA-5A/BR(A) HAS LYS-534 AND HPA-5B/BR(B)
CC HAS GLU-534. HPA-5B IS INVOLVED IN NEONATAL ALLOIMMUNE
CC THROMBOCYTOPENIA (NALT OR NATP). THE K534E POLYMORPHISM MAY PLAY A
CC ROLE IN CORONARY ARTERY DISEASE (CAD).
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD49b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17033; CAA34894.1;
DR EMBL; AF512556; AAM34795.1;
DR PIR; A33998; A33998.
DR PDB; 1A0X; 25-NOV-98.
DR PDB; 1DZ1; 02-AUG-01.
DR Genew; HGNC:6137; ITGA2.
DR MTM; 192974;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007596; F:blood coagulation; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;

3D-structure. 1 29
KW SIGNAL 30 1181
FT CHAIN 30 1181
FT DOMAIN 1133 1154
FT TRANSMEM 1155 1161
FT DOMAIN 1155 1161
FT DOMAIN 1155 1161
FT REPEAT 45 103
FT REPEAT 2 7
FT REPEAT 188 378
FT DOMAIN 378 433
FT REPEAT 434 486
FT REPEAT 488 549
FT REPEAT 551 610
FT REPEAT 615 667
FT CA_BIND 499 507
FT CA_BIND 563 571
FT CA_BIND 627 635
FT SITE 1157 1161
FT DISULFID 83 92
FT DISULFID 680 737
FT DISULFID 789 795
FT DISULFID 865 876
FT DISULFID 1019 1050
FT CARBOHYD 1055 1060
FT CARBOHYD 112 112
FT CARBOHYD 343 343
FT CARBOHYD 432 432
FT CARBOHYD 460 460
FT CARBOHYD 475 475
FT CARBOHYD 699 699
FT CARBOHYD 1057 1057
FT CARBOHYD 1074 1074
FT CARBOHYD 1081 1081
FT VARIANT 534 534
FT TURN 170 171
FT STRAND 173 180
FT TURN 183 184
FT HELIX 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT STRAND 220 224
FT TURN 226 228
FT HELIX 232 240
FT TURN 241 241
FT HELIX 252 262
FT TURN 263 264
FT HELIX 266 268
FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT HELIX 292 301
FT TURN 302 303
FT STRAND 304 311
FT HELIX 313 317
FT TURN 318 319
FT HELIX 323 330
FT TURN 331 332
FT HELIX 337 340
FT STRAND 341 344
FT HELIX 347 353
FT HELIX 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;
Query Match 18.2%; Score 1058; DB 1; Length 1181;
Best Local Similarity 26.9%; Pred. No. 3.le-64;
Matches 329; Conservative 214; Mismatches 494; Indels 188; Caps 43;
dbSNP:1801106;
/FTIG=VAR_003977.

INTEGRIN ALPHA-2.

EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

INTERACTION WITH HPS5.

FG-GAP 1.

FG-GAP 2.

VWFA.

FG-GAP 3.

FG-GAP 4.

FG-GAP 5.

FG-GAP 6.

FG-GAP 7.

POTENTIAL.

POTENTIAL.

GFFKR MOTIF.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

K -> E (IN ALLOANTIGEN HPA-5B;

dbSNP:1801106).

/FTIG=VAR_003977.

Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)
(CD49a).
ITGA1.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
MEDLINE=90338125; PubMed=2380249;
RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,
Each F., Carbonetto S., Reichardt L.F.;
"Molecular cloning of the rat integrin alpha 1-subunit: a receptor
for laminin and collagen."
J. Cell Biol. 111:709-720(1990).
[2]
X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.
MEDLINE=99313197; PubMed=1038626;
RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Kotliansky V.,
Gotwale P.J., Karpusas M.;
"Crystal structure of the alpha1beta1 integrin I-domain: insights into
integrin I-domain function."
FEBS Lett. 452:379-385(1999).
RL
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X52140; CRA36384.1;
DR PIR; A35854; A35854.
DR PDB; 1CK4; O3-MAY-00.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Repeat; Calcium; Magnesium; 3D-structure.
FT SIGNAL 1 28
FT CHAIN 29 1180
FT DOMAIN 29 1142
FT DOMAIN 1143 1165
FT TRANSMEM 1166 1180
FT DOMAIN 1166 1180
FT CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 103
FT REPEAT ? ?
FT REPEAT 175 388
FT DOMAIN 377 432
FT REPEAT 433 484
FT REPEAT 485 565
FT REPEAT 567 626
FT REPEAT 629 681
FT CA_BIND 497 505
POTENTIAL.

587 579 579 587
641 641 641 641
1172 1168 1172 1172
82 82 82 82
696 687 696 696
755 702 755 755
813 807 813 813
885 877 885 885
1062 1029 1062 1062
1073 1066 1073 1073
100 100 100 100
105 105 105 105
112 112 112 112
217 217 217 217
317 317 317 317
341 341 341 341
402 402 402 402
418 418 418 418
459 459 459 459
531 531 531 531
698 698 698 698
747 747 747 747
779 779 779 779
820 820 820 820
839 839 839 839
882 882 882 882
907 907 907 907
938 938 938 938
965 965 965 965
973 973 973 973
1007 1007 1007 1007
1084 1084 1084 1084
1103 1103 1103 1103
1114 1114 1114 1114
1180 AA; 130808 MW; 855DA2BE02362EE4 CRC64;
SQ SEQUENCE 1180 AA; 130808 MW; 855DA2BE02362EE4 CRC64;
Query Match 18.1%; Score 1062; DB 1; Length 1180;
Best Local Similarity 27.4%; Pred. No. 8e-64;
Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;
Qy 1 FNLDTENAMTFQENARG-FGQSVVQL---OGRVWVCAQOEIVAAVNOGRSLYQCDYGTGS 56
Db 29 FNVDRVNSMFSFGVEDMFGYTVQYQYENEGKWLIGSLVGPQKARTGDYKCPVGRER 88
Qy 57 CEP-IRLOVPVEA-----VNMSLGLSLAATTSPPQLLACQPTVHQTCSEYVYKGL 106
Db 89 AMPCVKLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNNGGFLACGLYVYRGCHLHYTTGI 147
Qy 107 CFLFGSNLRQOPKFPFALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLK- 165
Db 148 CSDVSPTFQVNSFAP--VQECSTQ-LD1VIVLDGNSIYP--WESVIAFLNDLLKRMDI 202
Qy 166 -KSKTLFSLMOYSEEFRIHFTKFEQNNPNRSLIKPITQLLG-RTHATGIRKVKVREL 223
Db 203 GPKQTQVGIQYGENVTHEFNLNKYSSTEEVLVAANKIGRGGGLQTMALGIDTARKEAF 262
Qy 224 NITNGARKNAFKILITDCEKFGDPLGYEDVLPADREGVIRVIGVGAFA-SE 278
Db 263 TEARGARGVKVVMVIVTDGESH-DNVRKQVIODCEDENIQRPISAILGHYRNGNSTE 321
Qy 279 KSRQELNVASKPPRPHVFOINNFEALKTQNLKRIKIFAIEGTQTQSSSSFFHEMSQEG 338
Db 322 KFVEEIKSIASEPTEKHFFNVSDDELALVTIVKALGERIFALEATADQSAASFEMESQTG 381
Qy 339 FSAAITSNGLPLSTVSGSYDWAGGVFLYTSKEKSTFFINMT--RVDSDND---AYLGAAA 393
Db 382 FSAHYSQDWMLGAVGAYDNWGTVMOKANQMVIPTNTTFTQTEPAKNEPLASLYGTVN 441
Qy 394 IILNRVQSLVGLGAPRYQHIGLVAMPFRONTGMESNANVKGTQIGAYFGSLCSVDVDSN 453
Db 442 SATIPGDLVLIAGOPRYNHTGVVYIKMEDGNINILQTLGGEQIGSYFGSVLTIDDKD 501
Qy 454 GSTDVLVLIGAPHY-----YEQTR--GGQVSVCPPLPRGQARWQCDAVLGE 497

Search completed: November 25, 2003, 14:17:34
Job time : 13.1742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:11:54 ; Search time 33.1068 Seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481b-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPPGAE PQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rhodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvrius:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4381	74.6	1151	11	Q9J130
2	3901.5	66.4	1036	11	Q8CA73
3	3809.5	64.8	920	6	Q28984
4	3480	59.2	1169	4	Q81VA6
5	3320.5	56.5	1169	11	Q9QXH4
6	3244.5	55.2	1161	11	Q9QYE7
7	1534.5	26.1	1161	11	Q9RTV4
8	1524	25.9	1160	11	Q9R200
9	1409	24.0	1196	13	Q9RTF1
10	1359.5	23.1	1086	4	Q96HB1
11	1358.5	23.1	1187	13	Q98TF0
12	1278	21.7	927	6	Q8HZV0
13	1167.5	19.9	1167	11	Q88340
14	1123	19.1	1167	11	Q88341
15	1053	17.9	1171	13	Q42094
16	1045	17.8	1038	11	Q8BS01

17	1018.5	17.3	1160	6	Q8MKF4
18	1018	17.3	895	11	Q9WUF8
19	865	14.7	348	4	Q8TES5
20	850	14.5	1332	5	Q9BPQ8
21	808	13.8	205	11	Q63001
22	753.5	12.8	780	13	Q06271
23	738	12.6	823	4	Q8WY18
24	686.5	11.7	823	11	Q8CE84
25	669	11.4	1032	11	Q61989
26	643	10.9	1036	11	Q91YD5
27	642.5	10.9	1033	6	Q9BGU3
28	623.5	10.6	257	11	Q8C270
29	619.5	10.5	1041	5	Q9UB90
30	614.5	10.5	1041	5	Q76378
31	579.5	9.9	1054	5	Q9U6S1
32	578.5	9.8	1034	13	Q98TT7
33	550	9.4	1033	13	Q42598
34	534	9.1	1016	13	Q91779
35	530	9.0	974	11	Q924W2
36	529	9.0	1073	11	Q8CC06
37	526	9.0	1047	6	Q9MZD6
38	525.5	8.9	1007	6	Q9GK48
39	512.5	8.7	1034	6	Q9TUN4
40	509.5	8.7	1036	6	Q9TUN6
41	506.5	8.6	1012	11	Q70304
42	491.5	8.4	1049	5	Q8Y51
43	475	8.1	833	5	Q9BPQ7
44	469.5	8.0	1036	6	Q9TU44
45	468.5	8.0	1034	11	Q8BQ25

ALIGNMENTS

RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.
 AC Q9J130;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Integrin beta 2 alpha subunit.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fathallah D.M. Sr., Zeria K. Jr.;
 RT "Cloning of the rat CD11b cDNA sequence."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF268593; AAF81280.1; -
 DR HSSP; P11215; 1BHQ.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 5_
 DR Pfam; PF00357; integrin_A; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VVFA; 1.
 SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.6%; Score 4381; DB 11; Length 1151;
 Best Local Similarity 72.7%; Pred. No. 1.6e-316;
 Matches 827; Conservative 150; Mismatches 158; Indels 2; Gaps 2;

Oy 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQRGSLQCQDYSGSCEPI 60
 |||||
 Db 17 FNLDTENPMTFQENASFGQSVIQLGETRVVVAAPQEVKAVNQIYALYQCDYSTNRCDPI 76
 |||||

```
QY 61 RLOVPVAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFGLSGNLRRQPOK 120
DB 77 PLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHQNCKNTYVYKGLCYFLSGNLRRKPOQ 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIISTVMEOLKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQOESNIAFLIDGSGSINTIDFQKMKFEFVSTVMDQFQSKTLFSLMOYSDFF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 240
DB 197 RTHFTFDFKRNPDKSHVRPIQLNGRTKTASGIRKVVRELFOKINGARDNAAKILV 256
QY 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDGDAFSEKSRQELNIVASKPPRDHVFQ 300
DB 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDGDAFSEKSRQELNIVASKPPRDHVFQ 316
QY 301 NFPAALNTIONQLREKIFALEGTQGTSSSFEHMSOEGFSAATITSGPLLSLVGSDWAG 360
DB 317 NFPAALNTIONQLREKIFALEGTQGTSSSFEHMSOEGFSAATITSGPLLSLVGSDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GAFLYPSKDKASPIINTRIDSDMNDAYLGYASAVISRNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQVSVCP 496
QY 481 PRGORARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 540
DB 497 PRGORARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 555
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQGHVLLRSQ 600
DB 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQGHVLLRSQ 615
QY 601 PVLVRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
DB 616 PVLVRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 675
QY 661 YDLALDSGRPHSRAVFNKTNTRTOTVGLGLTQTCETLKLQCPNIEDPVSIVLRNLF 720
DB 676 YDLALDSGRPHSRAVFNKTNTRTOTVGLGLTQTCETLKLQCPNIEDPVSIVLRNLF 735
QY 721 SLVGTPLSAGNLRPLVLAEDAQRLFTALFPFKNCGNDNICQDLSITFSFMSLDCLVWG 780
DB 736 SLVGTPLSAGNLRPLVLAEDAQRLFTALFPFKNCGNDNICQDLSITFSFMSLDCLVWG 795
QY 781 GPREFNVTVVRNDGDSYRTQVTFPEPLDLSYRKVSTLONORSQSRWLACASSTEV 840
DB 796 GPREFNVTVVRNDGDSYRTQVTFPEPLDLSYRKVSTLONORSQSRWLACASSTEV 854
QY 841 SGALKSTCSGINHPFIPENSEVFNTFFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 855 SGALKSTCSGINHPFIPENSEVFNTFFDVDSKASLGNKLLKANVTSENNMPTNKTEF 914
QY 901 QLELPVKAIVYVMTVSHGVTYKYNLTASNTSRVMOHQYOVNGLQSRSLPISLVFLVPV 960
DB 915 QLELPVKAIVYVMTVSHGVTYKYNLTASNTSRVMOHQYOVNGLQSRSLPISLVFLVPV 974
QY 961 RLCQTVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAIPVNCISIAVCORIQDIP 1020
DB 975 RLCQTVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAIPVNCISIAVCORIQDIP 1034
QY 1021 FFGIOEBFNATLKNGLSFDWYIKTSHNHLIVSTAEILFNDVSTFLLPGOGAFVRSOTET 1080
DB 1035 FFGIOEBFNATLKNGLSFDWYIKTSHNHLIVSTAEILFNDVSTFLLPGOGAFVRSOTET 1094
QY 1081 KVEPFEPNPLPLTVGSSVGLLLALITAAALYKGLFFKRYQKMDMSEGGPPGAEPQ 1137
DB 1095 KVEPFEPNPLPLTVGSSVGLLLALITAAALYKGLFFKRYQKMDMSEGGPPGAEPQ 1151
```

```
RESULT 2
QBCA73 ID Q8CA73 PRELIMINARY; PRT; 1036 AA.
AC Q8CA73;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Integrin alpha M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK039444; BAC30350.1; --
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;

Query Match 66.4%; Score 3901.5; DB 11; Length 1036;
Best Local Similarity 66.1%; Pred. No. 6.1e-281;
Matches 752; Conservative 125; Mismatches 142; Indels 119; Gaps 2;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGRVYVVGAPQEIIVAAHQRSGLYQCDYSTGSCPEI 60
DB 17 FNLDTEHPMTFOENAKFGQSVVQLGGTSVVVAAVQAEKAVNQTGALYQCDYSTSRCHPI 76
QY 61 RLOVPVAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFGLSGNLRRQPOK 120
DB 77 PLQVPVAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKNTYVYKGLCYFLSGNLRRPQO 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIISTVMEOLKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQOESNIAFLIDGSGSINTIDFQKMKFEFVSTVMDQFQSKTLFSLMOYSDFF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 240
DB 197 RHFTFDFKRNPDKSHVRPIQLNGRTKTASGIRKVVRELFOKINGARDNAAKILV 256
QY 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDGDAFSEKSRQELNIVASKPPRDHVFQ 300
DB 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDGDAFSEKSRQELNIVASKPPRDHVFQ 316
QY 301 NFPAALNTIONQLREKIFALEGTQGTSSSFEHMSOEGFSAATITSGPLLSLVGSDWAG 360
DB 317 NFPAALNTIONQLREKIFALEGTQGTSSSFEHMSOEGFSAATITSGPLLSLVGSDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GAFLYTSKDKVTFINTRVDSMDNDAYLGYASAVISRNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGQVSVCP 480
DB 437 ENFGTWEPTISNG----- 450
QY 481 PRGORARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVAIGAPCEDNRGAVYLF 540
DB 451 ----- 450
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQGHVLLRSQ 600
DB 451 -----SORIIGAHPSPGLQYFGQSLSGGQDLTMDGLMDLAVGAQGHVLLRAQ 498
QY 601 PVLVRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
DB 499 PVLVRVKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGQIOSVVT 558
```

```
QY 661 YDLALDSGRPHSRVAFNETHKSTRQTOVLGLTQTCETLKLQNCIBDPSPVILRLNF 720
DB 559 YDLALDPCRIRAFDETKNTRRTQVFLGMQKCELTKLJLPDCVDDSVPIILRLNY 618
QY 721 SLVGTPLSARGLRPLVLAEDQRPLTALFPEKKNCGNDNICODDLSITFSFMSDCLVWG 780
DB 619 TLVGEPLASFGNLRPLVLAEDQRPLTALFPEKKNCGNDNICODDLSITFSFMSDCLVWG 678
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRORSWRL-ACSSASSTE 839
DB 679 GPDEFNMSVTLRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRORSWRL-ACSSASSTE 738
QY 840 VSGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTE 899
DB 739 GHGALKSTTNINHPFIPANSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTE 798
QY 900 FOELPLPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVLVP 959
DB 799 FOELPLPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVLVP 858
QY 960 VRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDI 1019
DB 859 VOINNVTVMHDPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDI 918
QY 1020 PFFGIQEBFNATLKNLSFDMYIKTSHNHLIVSTAEILFNDSTFTLLPGQCAFVRSQTE 1079
DB 919 PSFNTQEIFNVTLKNLSFDMYIKTSHNHLIVSTAEILFNDSTFTLLPGQCAFVRSQTE 978
QY 1080 TKVEPFEVNPBLPLIVGSSVGLLALLIITAAALYKLGFFKQYQKDMMEGPPGAPQ 1137
DB 979 TKVEPFEVNPBLPLIVGSSVGLLALLIITAAALYKLGFFKQYQKDMMEGPPGAPQ 1036

RESULT 3
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CD11b (Fragment).
GN CD11b.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AA816869.1; -.
DR HSP; F11215; I8HQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.8%; Score 3809.5; DB 6; Length 920;
Best Local Similarity 78.9%; Pred. No. 3.5e-274;
Matches 727; Conservative 84; Mismatches 109; Indels 1; Gaps 1;

QY 118 POKPEALRGCPQESDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKKSKTLFSLMYS 177
DB 1 POKPEALRGCPQESDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKKSKTLFSLMYS 60
```

```
QY 178 EEFRIHTFFKEFQNNPNRSLIKPITQLLQRTHTATGIRKVRLELNTNGARKNAFKIL 237
DB 61 EDFYTHFTFDKRNPSKLLVRPIRQLLQRTHTATGIRKVRLELNTNGARKNAFKIL 120
QY 238 ILITDGSKFGDPLGYEDVPEADREGVIRVIVGVDAFRSEKSRQELNTVASKPRDHVF 297
DB 121 WITDGSKFGDPLGYEDVPEADREGVIRVIVGVDAFRSEKSRQELNTVASKPRDHVF 180
QY 298 QINNPEALKTIQNLREKIFAIEGTQTQSSSSPHEMSQGFSAAITNSGPLLSTVGSYD 357
DB 181 QVNNFEAVKTIQNLREKIFAIEGTQTQSSSSPHEMSQGFSAAITNSGPLLSTVGSYD 240
QY 358 WAGGVFLYTSKEKSTFINMTRVDSMDMDAYLGYAAAIILNRNVQSLVGLGAPRYQHILVA 417
DB 241 WAGGAFLHMPKDRVIFINTTRVDSMDMDAYLGYAAAIILNRNVQSLVGLGAPRYQHILVA 300
QY 418 MFRONTGWESNANVKGTQIAYFGASCLSVVDVDSNGSTDLVLIGAPHYEQRTRGGQSV 477
DB 301 MFKONGAWEKNAIDIKGSIQSYFGASCLSVVDVDSNGSTDLVLIGAPHYEQRTRGGQSV 360
QY 478 CPLPRGORARWQCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAV 537
DB 361 CPLPQG-RAKWCQKRVILCGQGHPSRFGAALTALGDVNGDKLTDVAIGAPGEDNRGAV 419
QY 538 YLFHGTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVDLTVGAQGHVLL 597
DB 420 YLFHGTSELGISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVDLTVGAQGHVLL 479
QY 598 RSQPVLRKAIEMFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQS 657
DB 480 RSQPVLRKAIEMFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQS 539
QY 658 VVTYDLALDSGRPHSRVAFNETHKSTRQTOVLGLTQTCETLKLQNCIBDPSPVILRL 717
DB 540 ITTYDLALDPCRPHPRAVFEETKNTRRQTGLSKRCEHLALWLPDCVEDSVTPVILR 599
QY 718 INFSLVGTPLSARGLRPLVLAEDQRPLTALFPEKKNCGNDNICODDLSITFSFMSDCL 777
DB 600 INFSLVGTPLSARGLRPLVLAEDQRPLTALFPEKKNCGNDNICODDLSITFSFMSDCL 659
QY 778 VVGPRFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRORSWRLACSSASSTE 837
DB 660 VVGPRFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRORSWRLACSSASSTE 719
QY 838 TEVSGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNK 897
DB 720 TEVSGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNK 779
QY 898 TEFOLELPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVL 957
DB 780 TEFOLELPVKYAVTVMTSHGVSTKYLNFTASENTSRVMOHOQVSNLQORSPLISLVL 839
QY 958 VPVRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQ 1017
DB 840 VPVRLNQTIVDRPQVTFESNLSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQ 899
QY 1018 DIPFGIOEBENATLKNLSF 1038
DB 900 DIPFGIOEBENATLKNLSF 920

RESULT 4
Q81VA6 PRELIMINARY; PRT; 1169 AA.
ID Q81VA6
AC Q81VA6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin, alpha X (Antigen CD11c (p150), alpha polypeptide).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Blood;
RA  Strausberg R.;
RL  Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC038237; AAH38237.1; -
SQ  SEQUENCE 1169 AA; 128521 MW; 17B484FEFC79EB6 CRC64;

Query Match      59.2%; Score 3480; DB 4; Length 1169;
Best Local Similarity 61.2%; Pred. No. 1.7e-249;
Matches 691; Conservative 141; Mismatches 291; Indels 6; Gaps 4;

QY  1  FNLDTENAMTFOENARGCQSVVQLQSGSRVVGAPQEIIVAAQNRGSLYOCDSYSCCEPI 60
DB  20  FNLDTEELTAFRVDSAGFSDVQYANVSWVVGAPQKITAANQQTGLYQCGYSTGACEPI 79

QY  61  RLQVPEAVNMVSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLGSLNLRQPOK 120
DB  80  GLQVPEAVNMVSLGLSLAATTSPQLLACGPTVHHECGRMVYLTGLCFLGLPT--QLTQR 137

QY  121  FPEALRGCPQEDSIAFLVDGSGSIIPHDFRAKEFISTVMQKSKTLFSLMQYSEEF 180
DB  138  LPVSRQCEPRQEDIVFLIDGSGSISSRNFAFMNEFRAVISQFORPSTQFSLMOPSNKF 197

QY  181  RIHFTKBPQNNPNSLIKPIITQLGRTHATGIRKVVRELPNITNGARKNAFKLILI 240
DB  198  QTHFTFEFRSSNPLSLASVHQLQGYTATAIQNVVHRUFPHASYGARRDAAKLIVI 257

QY  241  TDCEKFGDPLGYEDVPEADREGVIRYVGVGDAFRSEKSRQELNTVASKPRDRHVFQIN 300
DB  258  TDCKKSGDSLVDKVDVPMADAAGIIRYAGVGLAFQNRNSKELNDIASKPSQEHIFKVE 317

QY  301  NFPAKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAIITSNPLLSYVDWAG 360
DB  318  DFDALQIQNLREKIFAIEGTETTSSSFELEMAQEGFSAVTPDGPVLGAVGSPWMSG 377

QY  361  GVFLYTSKSTSTINTRVDSNMNDAYLGAAAILLRNVQSLVLCAPRYOHLGVAMFR 420
DB  378  GAFLYPNMSPTEINNSQENVDRDYSVLGYSTELALWQVSLVLCAPRYOHTGKAVIF 437

QY  421  QNTGMNESANVAGTQIGAFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCPL 480
DB  438  QVSRQWRMAEVTGTQIGSYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCPL 497

QY  481  PRGORARWCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB  498  PRGWR-RWVCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVVI GAPGEKENRGAVYLF 556

QY  541  HGTSGSGISPSHSQIRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB  557  HGVLPISPSHSQIRAGSKLSPRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616

QY  601  PVLVRKAIMFENPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDRLRGQIQSVVT 660
DB  617  PVLWVGVMQFIIPAEIPRSFAFECREQVVEQTLVQSNICLYIDKRNKLLGSLDQSSVT 676

QY  661  YDIALDSGRPHSRAVFNETKXSTRQVILGLTQTCTETLKLQLPNCIEDPVSPVLRILNF 720
DB  677  LDIALDPGRLSPRATFQETKXSTRVRLVGLKAHCENFNLLPSCVEDSVTEITLRILNF 736

QY  721  SLVGTPLSAGNLRPVLAEDAORLFTALFPFEKNCNDNTCODDLSITFSFMSLCLVVG 780
DB  737  TLVKGKPLAFNRNPLMLAADAQRYFTASLPFEKNCAGADHICQDNLGISFSFPGKLLVVG 796

QY  781  GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACBSASSTEV 840
DB  797  SNLELNAEVMVWNDGDSYGTVTFSHPAGLSYRYVAEGQKQQLSLHLTCDSPVGV-- 854

QY  841  SGALKSTSCINHPIFPENSEVFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTFF 900
DB  855  SQGTWSTSCINHLIPFGGAQIITFLATFDVSPRAVLGDRLLLTANVSENNTPTSTKTF 914

QY  901  QLELPVKYAVYVTVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
DB  960  VRLNQTVIWRDPQVTFSENLSSTCHTHERLPSSHDSFLAELKAPVNCISIAVCORIQCDI 1019
DB  975  VELNQEAVMMDEVSHFQNPSPSRCSSEKIAPPASDFLAHQINPVLDCSIAGCLRFRCDV 1034
QY  1020  PFFGIQEEFNATLKNLSFDWYIKTSHNHLILIVTAETILFNDVSFTLLPGGQAFVRSQTE 1079
DB  1035  PSFSVQEEELDTLTKNLSFGWVRQLQKXSVSVSAEITFTSVYSQLPQGEAFMRAQTT 1094
QY  1080  TKVPEFPEVNPPLIVGVSSVGLLLALITAAALYKLGFFKRYQKDMSE 1128
DB  1095  TVLEKYKVHNPPLIVGVSSIGLLALLITAVLYKVGFFKRYKEMMEE 1143

RESULT 5
QYQX4
ID  Q9QX4; PRELIMINARY; PRT; 1169 AA.
AC  Q9QX4;
DT  01-MAY-2000 (TRENBLrel. 13, Created)
DT  01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT  01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE  Leukocyte adhesion glycoprotein p150.95 alpha integrin subunit.
GN  ITGAX.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA  Tsuchiya H.;
RT  "Isolation of Genes Selectively Expressed by Dendritic Cells.";
RL  Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF211864; AAF23492.1; -
DR  HSSP; P11215; 1BHQ.
DR  MGD; MGI:96609; Itgax.
DR  InterPro; IPR000413; Integrin_alpha.
DR  InterPro; IPR002035; VWFA_A.
DR  Pfam; PF01839; FG-GAP; 5.
DR  Pfam; PF00357; Integrin_A; 1.
DR  Pfam; PF00092; vwa; 1.
DR  PRINTS; PR01185; INTEGRINA.
DR  PRINTS; PR00453; VWFADOMAIN.
DR  SMART; SM00191; Int_alpha; 4.
DR  SMART; SM00327; VWA; 1.
DR  PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR  PROSITE; PS50234; VWFA; 1.
DR  Integrin.
KW  SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match      56.5%; Score 3320.5; DB 11; Length 1169;
Best Local Similarity 56.8%; Pred. No. 1.3e-237;
Matches 647; Conservative 172; Mismatches 302; Indels 19; Gaps 7;

QY  1  FNLDTENAMTFOENARGCQSVVQLQSGSRVVGAPQEIIVAAQNRGSLYOCDSYSCCEPI 60
DB  20  FNLDTEELTAFRVDSAGFSDVQYANVSWVVGAPKIKATNQIGLYKCGYHTGCEPI 79

QY  61  RLQVPEAVNMVSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLGSLNLRQPOK 120
DB  80  GLQVPEAVNMVSLGLSLAATTSPQLLACGPTVHHECGRMVYLTGLCFLGLPT--QLTQR 138

QY  121  FPEALRGCPQEDSIAFLVDGSGSIIPHDFRAKEFISTVMQKSKTLFSLMQYSEEF 180
DB  139  FTAQOCECPQODIVFLIDGSGSISSDTDFERKMLDFVKAVMSQLQRPSTFSLMQPSDYF 198

QY  181  RIHFTKBPQNNPNSLIKPIITQLGRTHATGIRKVVRELPNITNGARKNAFKLILI 240
DB  199  RVHFTFNFNISTSSPLSLGSRVRLRGYTYTASAKHVITELFTTQSGARQDATKVLIVI 258

QY  241  TDCEKFGDPLGYEDVPEADREGVIRYVGVGDAFRSEKSRQELNTVASKPRDRHVFQIN 300
```

```
Db 259 TDGRKQGNLSYDSVPWAEASIRYAIGVKAPYNEHSHKQELKATASMPSEHYVSFE 318
Qy 301 NFEALKTIQNLREKI FAIEGTOTGSSSSSFEHMSQEGFSAATNSGPLLSTVGSYDWAG 360
Db 319 NFDALKDLENQKKEI FAIEGTETPSSSTFEHMSQEGFSAVFPDGPVLGAVGSFWSG 378
Qy 361 GVFLYSKEKSTFIMNTVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 379 GAFLYPSNMWRPTFFINMSQENEDMRDYLGYSTALAFMGVHSLILGAPRHQHTGKWIPT 438
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQVSCPL 480
Db 439 QESRHWPKSEVRGTQIGSYFGASLCSVDMDRDSGLVLIGVPHYEHTTGGQVSCPM 498
Qy 481 PRGQARWQCDALVYGEQOPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 499 P-GVGRWHCGTTLGEGQHPWGRFGAALTVLGVDNGDSLADVAIGAPGEENRGAVYIF 557
Qy 541 HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 558 HGASRQDIAPSPQRISASQIPSRIOYFGQSLSGGQDLTRDGLVDLAVGSKRVLRLR 617
Qy 601 PVLVRKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVT 660
Db 618 PILRVSPVTHFTPABISRSVFECQVAPQTLSDATVCLHIHESPKTQL--GDLRSTVT 675
Qy 661 YDLALDSGRPHSRVFNETHKSTRQTVGLGTOTCETLKLQPNICIEDPSPVTLRLNF 720
Db 676 FDALDHGRLSTRAIFKETKTRALTRVKTGLNKHCBESVKULLPACVEDSVPTITLRLNF 735
Qy 721 SLVGTPLSAFGLRVLAEADQRLFTALPFEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGVPISSQLNLQPLAVDDQTYFTASLPFEKKNCGADHICODDLSVVFPGPDLKTLVG 795
Qy 781 GPRENVTVTVRNDGEDSYRTQVTPFFPDLDSYRKVSTLQ-----NORSQRSWR 829
Db 796 SDLELVNTVTVRNDGEDSYRTVTLFYPVGLSFRFAEGQVFLRKEDQOQRRGQSHLH 855
Qy 830 LACESASTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSE 889
Db 856 LMCD--STPDRSQGLWSTSCSRHVFIRGGQMTPLVTFVSPKAEGLDRLLRLRARGSE 913
Qy 890 NMMPRTNKEFOLELPVKYAVVMVTSHGVSPTKYNLFTASE-NTSRVNMHQVQVSNLQOR 948
Db 914 NNVPCTKTTLQLELPVKYAVYTMISSHDQFTKYNLFTSEKETSUVVEHFRQVNNLQOR 973
Qy 949 SLPISLVLVPLVRLNQTIVWRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNC 1008
Db 974 DVPVSINFWPIELKGEAVW-TVMVSHFPQNPLTQCYRNRLKPTQPDLLTHMQKSPVLDCS 1032
Qy 1009 TAVCQRIODIPFFGQIOREFNATLKNLSFDWYIKTSHNHLIIVSTAILENDSVFTLLP 1068
Db 1033 TADCLHLRCDIPSLGILDELVPILKGNISFGMISQTLQKQKLLSEAEITNTSVYSQLP 1092
Qy 1069 GOGAFVRSOTETKVPFFVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSE 1128
Db 1093 GQEAFLRAQTKTVLEMYKVNHPVPLIVGSSVGGLLLLAIITAILYKAGFFKRYQKEMLEE 1152

RESULT 6
Q9QYE7
ID Q9QYE7 PRELIMINARY; PRT; 1161 AA.
AC Q9QYE7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21341.1; -.
DR HSP; F11215; IBHQ.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.2%; Score 3244.5; DB 11; Length 1161;
Best Local Similarity 57.6%; Pred. No. 5.7e-232;
Matches 649; Conservative 163; Mismatches 302; Indels 13; Gaps 9;

Qy 2 NLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEI VAAANQSGSLYQCDYSTGSCCEPIR 61
Db 21 NLDVEEPIVREDAAASFGQIVVQGGSLVVGAPLEAVAVNQTGRLDCAPATGMCQPIV 80
Qy 62 LQVPEAVNNSLGLSLAATSPOLLACGPTVTHQTCSENTYVVKGLCFGLFSGNLRRQPKF 121
Db 81 LRSPLAVNNSLGLSLVTATNNAQLLACGPTAQRACVKNYAKGSCLLGSSL-QFIQAV 139
Qy 122 PEALRGCPQSDSDIAFLVDCSGSIIPHDFRRAKEFI STVMEQLKSKSTLPSLMQYSEFR 181
Db 140 PASMPECPROMEDIAFLIDGSGSINQRDFAQMDKFVKALMGFEFASTSTLSLMQYSILK 199
Qy 182 IHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVVRELFTNGARKNAFKILILIT 241
Db 200 THFTTFEKNILDPQSLVDPIVQLQGLTYTATGIRTVMBELFHSKNGSRKSAKILLVIT 259
Qy 242 DGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQELMTVAKPPDRHVFOINN 301
Db 260 DGQKYRDPLESDYDIPADKAGIIRYAGVDAFQEP TALKELNTIGSAPPQDHVFKVGN 319
Qy 302 FEALKTQNLREKI FAIEGTOTGSSSSSFEHMSQEGFSAATNSGPLLSTVGSYDWAG 361
Db 320 FAALRSIQRLQEKI FAIEGTQSRSSSFQHEMSQEGFSSALTS DGPVLGAVGSFWSG 379
Qy 362 VFLYTSKEKSTFIMNTVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 421
Db 380 AFLYPNTRPTFFINMSQENEDMRDYLGYSTALAFMGVHSLILGAPRHQHTGKWIPTQ 439
Qy 422 NTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGQVSCPL 481
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSGLVLIGAPHYYEQTGGQVSCPL 499
Qy 482 RQBARWQCDALVYGEQOPWGRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 541
Db 500 -GVGRWQCEATLHGEQHPWGRFGVALTVLGDVNGDNLADVAIGAPGEESRGAVYIF 558
Qy 542 GTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSOP 601
Db 559 GASRLIMPSPQRISASQIPSRIOYFGQSLSGGQDLTQDGLVDLAVGSKRVLRLRSLP 618
Qy 602 VLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVV 661
Db 619 LKVELSIRAPMEVAKAVVQCWERTPTVLEAGATVCLTVHKGSPDLL--GNVQGSVRY 676
Qy 662 DLALDSGRPHSRVFNETHKSTRQTVGLGTOTCETLKLQPNICIEDPSPVTLRLNFS 721
Db 677 DLDLDPGRILSRAIFDETCKNCTLGRKTLGLGDHCHCETVKULLPDCVEDAVSPIILRNFS 736
```

```
QY 722 LVGTPLSAFNGRNPVLADDAQRLFTALPFPEKNGCNDNICODDLSITSPMSLDCLVVG 781
Db 737 LVKDSASP-RNLHPVLAVGSDHIITASLPFEKCKCKELLCEGLGIFSNGLQVLVVG 795
QY 782 PREPNVTYVRNDGSDSYRTQVTFPPDLDSYRKVSTLQNSORSORSLACESASSTVS 841
Db 796 SPELTVTVTVMNEGSDSYGLVKFYFAGLSYRRVTGTQ-QPHQVPLRLACEAPAAQED 854
QY 842 GALKSTSCSINHPIIPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTBFQ 901
Db 855 --LRSSCSINHPIIFREGAKTTFMIFDVSYKAFGLDRLLRKAKASNNKPDNTKTAQ 912
QY 902 LELPKYAVYVMVTSYKLVNFTASENTR-VNQHQYQVSNLQORSLSPLSLVFLVPV 960
Db 913 LELPKYVTVYTLISRQEDSTNHFSSHGGRQEAHRYRNNLSPLKLAVRNVFVFPV 972
QY 961 RLNQTVINDRPOVTFSENLS--TCHTKERLPSHDSFLAELRKAPVNVCSIAVCQRIQCDI 1019
Db 973 LLNGVAVMD---VTLSPPAQGVSCVSKMPQNPDPFLQIQRRSVLDSCIADCLHFRCDI 1029
QY 1020 PFFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFLVRSQTE 1079
Db 1030 PSLDIODELDFILRGNLSFGWSQTLQEKVLLVSEAITFTDTSVYSQLPGQEAFLRAQVE 1089
QY 1080 TKVEPEVNPPLVGVSSVGGLLLALITALYKLGFPKQYKDM 1126
Db 1090 TTEELYVYVEPIFLVAGSVGGLLLALITVLYKLGFPKQYKEML 1136

RESULT 7
Q9WTV4 PRELIMINARY; PRT; 1161 AA.
AC Q9WTV4;
DT 01-NOV-1999 (TrEMBLrel. 12; Created)
DT 01-NOV-1999 (TrEMBLrel. 12; Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21; Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=DBA/2J; TISSUE=Spleen;
RA Ma R. Z.; Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; -.
DR HSSP; P20701; 1LFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWF_A; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;

Query Match 26.1%; Score 1534.5; DB 11; Length 1161;
Best Local Similarity 34.0%; Pred. No. 9.6e-105;
Matches 399; Conservative 217; Mismatches 456; Indels 101; Gaps 37;

QY 1 FNLDTENAMTFOENA-RGFQCSVVVQLGSRVVGAGPQEIIVAAHQRSGLYQCDYSTGSCPE 59
Db 24 YNLDRPTQSLAQAGRHFGYQVLQIEDG-VVVGAPCE---GDNTGLYHCRITSSEFCQP 79
QY 60 IRLQVPVEAVNMSLGUSLAATTSPQLLACGPTVHTQCSNTVYKGLCLFLGSLNLRQOPQ 119
```

```
Db 80 VSLH-GSNHTSKYLGWTLATDAAGSLACDGLSRTCDQNTVLSGLCYLFPQSLGPM 138
QY 120 KFPBALRGCOESDIAFLVDGSGSIIPHFRAKEFISTVMQKSKTLFSLMQVSEE 179
Db 139 QNRPAYQECMKGVKVDLVFLFDGSGSLDRKDFEILFEMKDMVRKLSNTSYQFAAVQSTD 198
QY 180 FRIHFTPEF-QNNPNPERSLIPKLTOLLGRTHATGIRKVVRELFNITNGARKNAFKILI 238
Db 199 CRTEFTFLDVKKQKNDVLLGVSQPMFLTNTFRALNVVAVHFKESGARPDATKVLV 258
QY 239 LITDGEKFGDPLGYEDVIEADREG-----VIRYVIGVDADFRSKSKQELNVTASKP 291
Db 259 IITDG-----EASDKGNISAAHDIRYIIGICKHFVSQVKOKTLHFASEP 304
QY 292 PRDIVFOINNPEAKTTONQLREKIFAIEGTQTGSSSSFEHMSQEQFSAAITNSGPLLS 351
Db 305 VEVEFKLIDTPEKLDKDLTDLQRIYAIETGNRODLTFSFMELSSSSISADLSKGHAVG 364
QY 352 TVGSYDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLVGAAA-IILNRNVQSLVLGAPR 409
Db 365 AVGAKDWAGGFLDLREDLQATFVQEPFLTSVDRGGYLGTVVAMTSSRSPRLAAAPR 424
QY 410 YQHIGLVAMF--QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYY 467
Db 425 YQHVQVLLFOAPEAGGRWNTQKIETQIGSYEGGELCSVDLDQDGEAELLIGAPLFF 484
QY 468 BOTRGQGVVCPPLPRGORARWQCDVLYGEGQWPGAFGAALTVLGVNGDKLTDVAIGA 527
Db 485 GEORGRGVFTY---QRRQSLFEMVSELQGDPGYPLGRFGAAITALTIDINGDRLTDVAVGA 541
QY 528 PGEEDNGAVYLFHGTSGSISPSHSORISAGSKLSPLQYFGOSLSGGODLTMDGLVPLT 587
Db 542 PLEE--QGVYIFNGKPG-GUSPQPSRIQQAQVFPGRWFGRIHGVKDLGGDBLADV 598
QY 588 VQAQHVLLLRSPQLRVKAIEMFNPREVARNVFECNDQVQVKGKAG-EVRVCLHVQKST 646
Db 599 VGPEGRVVVLSRRPVDVVTLSFSPEIIPVHEVECSYSAREEQHVKLACFRKIKPLT 658
QY 647 RDRUREGQISVVYDIALDGRPHSRVAFNETKNSTRQTVGLTLTCTETLKLQLPNC 706
Db 659 PQ--FQGRLLANLSYTLQLDGHRMRSGLFPDGGSHELSGNTSITP-DRSCLDLHFHFPIC 715
QY 707 IEDPVSPIVLRNLSLV---GTPLSAFCN-LRPVLAEDAQRLFTALPFPEKNGCNDNICQ 762
Db 716 IODLISPINVSFLNLSLEEETPRDQKRAMQPIRLPSIHV-TKEIPFNKCGEDKKCE 774
QY 763 DDLSTIFSFMSLDCLVWGGP-----REFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKV 816
Db 775 ANLTLSPPARS-----GPLRLMSSASLAVEMWTLNSGDAYVWRLDLDPPRGLSPRKV 827
QY 817 STLQNRORSORWRLACESASSTEVSGAL-KSTSCSINHPIIPENSEVFNITFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCEEL--TEGSSLLTKLTKCNVSSPIFKAGQEVSLQVMFNTLLSS 882
QY 876 LGNKLKLLKANVTSEN-NMPTNKTQFQLELPVKYAVYVMVTSYKLVNFTASENTR 934
Db 883 WEDFVFLNGTVHCENENSSLDQEDNSAATHIPVLYPNILTKEQENSTLYISTPKGPKTQ 942
QY 935 VMQHQQV---SNLQORSLSPLVFLVPVRLNQTVIWRPQ--VTFSENLS--TCHT 985
Db 943 QVQHYVQVRIQPSAYDNKMP-TLEALVGPV-----WPHSEDPITYTWSVQTDPLVTCHS 995
QY 986 KE-RLPSSHDFLAELRLKAPVNVCSIAVCQRIQCDIPFPGIQEEFNATLKGNSLFDWYIKT 1044
Db 996 EDLKRPSSE--AEQPCPLPGV-----QFRCPIVF---RREILIQVTGTVELSKEIKA 1041
QY 1045 SHNHLIVSTAEILFNDSVFTLLPGQGAFLVRSQTEKVEFVFPNPPLVLCVSSVGGLL 1104
Db 1042 S-STLSLCSLSVSNSSKHFHLYGSKA-SEAQVLVKVDLIHEKMLHVVLVSGGLVL 1099
QY 1105 LALITAAALKGFFKQYKDMV-SEGGPPGAEP 1136
```


[illegible]

Db 118 SYKTEFDSYVKKRDPDALLKHVKHMLLLTNTFTGAINVYVATEVFEELGARDPDKVLI 177
Qy 239 LITDGEFGPLGVEDVPEADREGVIRYVIGVGDAPRSEKSEKQELNTVASKPRDHVFO 298
Db 178 IITDGE--ATDSGNIDAOK-----IIRYIIGIKGHFQTKESQETLHKFASKASFEVKI 230
Qy 299 INNEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFAAITSNGPLSTVGSYDW 358
Db 231 LDITFEKLKDLFTELQKIVIEGTSKQDLTSFNWELSSGISADLSRGAHVAVGAKOW 290
Qy 359 AGGVF-LYTSKEKSTFINMTRVSDMDNDAYLGAAA-IILNRVQSLVLCAPRYQHIGLV 416
Db 291 AGGFLDLKADLQDDTFIGNEPLPEVRAGVLYVTWLPSPRQKTSLLASGAPRYQHMGV 350
Qy 417 AMFR-QNTGWESNANVKGTGTGAYEGASICSVDVDSNGSTDLVLIGAGHYEQRGGQ 474
Db 351 LLFQEPQGGHWSQVTHGTQSGYFGBELCGVDQDGETELLIGAPLFYGEORGR 410
Qy 475 VSVCLPRGRARWQCDV--LYGEOQPMGRGAALTVLGDVNGDKLTDVAICAPGEED 532
Db 411 VFYI-----QRQLGFEVSELOQDPCYPLGRFGEAITALTDINGDLVDVAVGAPLEE- 464
Qy 533 NRGAVYLFHTSGSGISPSHSQRTAGSKLSPRIQYFGOSLSGGODLTMDGLVLTVAQG 592
Db 465 -QGAIVYFNGRHG-GLSPQSPRIEGTVQLSGIQWFGRSIHGVKDLGDLGADVAVGAE 522
Qy 593 HVLLRSQPLRVKAINFENPREVARNVFECNDVV-KGEAGEVRVCLHVQSTRDLR 651
Db 523 QMIVLSRPVDMVTLSFSPAETPVHEVECSYSTSNMKEGVNITICQI-KSLIPQF- 580
Qy 652 EGOJOSVVTVDALDGRPHSAVFNETKSTROTGLTGTOTCELTKLQPNCTEDPV 711
Db 581 QGRIVANLTTLQDGHTRRRGLFPGRHELNRNIAT-TSMSCDTDFSHFPVCVDLLI 639
Qy 712 SPIVLRNLSL---VGTPLSAFGN-----LRPVLARDAQRLEFALPPEKNCNQNICQ 762
Db 640 SPINVSLSFWESEGTPRORAGKDIPILRPSLHSETWEI-----PFEKNCEDKCE 694
Qy 763 DDLSTISFMSLCLVGGPREFNVTVNRDGEDSVRTQVTFPPLDLSYRKVSTLQNG 822
Db 695 ANLRVSFSPARSALRLTAFASLSVELSLNLEEDAYVQDLHFPPLGSLFRKVEML--- 751
Qy 823 RSQSRWELACES--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKL 880
Db 752 KPHSQIPVSCLEELPEERLSRAL---SCNVSSPIFKAGHSVALQMFMTLVNWSGDSV 808
Qy 881 LLKANVTSENN---MPTNKTEFQLELPVKYAVYVVTSHGVSTKYLNFTASENTSRV 936
Db 809 ELHANVTNNEDSDLEDNSATTI---IPILYPINILIQOEDSTLVVSTPKGPKHQV 865
Qy 937 QHOYQV---SNLQORSPL-ISLFLVPLVRLNQTVWRDPOVTFSENLSSTCHTK--ERLP 990
Db 866 KMYQVRIQPSIHDHNIPTLEAVVGVPPSEGPITHQVSVQMEPPV--PCHVEDLERLP 923
Qy 991 SHSD--FLAELKAPVNCISIAVCORIQCDIPFGIOEEFNATLKNLSFDWIKTSHNH 1048
Db 924 DAAEPCPLGALFRCPPV-----FRQETLVQVIGTLELVGSEIAS-SM 964
Qy 1049 LLIVSTAEILFNDSVFTLLPQGAFAVRSQETKVEPEVFNPLPLIVSSVGGLLILLI 1108
Db 965 FSLCSSLISFNSKPHLYGSNASL-AQVVMKVDVYVQKMLYLYVLSGIGLLILLI 1023
Qy 1109 TAALYKLGFTFKROYKQMMSEG-GPPGAEP 1136
Db 1024 FIVLYKVGFFKRNKERNKEAGRGVPGIP 1052

RESULT 11
Q98TFO
ID Q98TFO PRELIMINARY; PRT; 1187 AA.
AC Q98TFO;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DB CD11-2.
GN CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB048537; BAB39135.1;
DR HSSP; P20701; ILPA.
DR InterPro; IPR001969; Aspartate site.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;
Query Match 23.1%; Score 1358.5; DB 13; Length 1187;
Best Local Similarity 31.6%; Pred. No. 1.2e-91;
Matches 377; Conservative 206; Mismatches 452; Indels 157; Gaps 43;
Qy 1 FNLDTENAMTFOENARG-FGQSVVQLQ-GSR--VVVGAPOEIVAAQRGSLYQCYSTGS 56
Db 32 FNIDTSHPLRNFCTPEDFFGYSVYQTEFGNRKQIIVGAPLE---GNSAGEMYSTADLQS 88
Qy 57 CEPRIQLQV-----VEAVNMSLGLSLAATTSPOLLACGTVTHTCSTYVKGCLFLFGS 112
Db 89 CK--RLQRFSGESVRFEGMSAAVSSAALTS-----CSPYFAHECDGNSYLVNGCYQFNS 140
Qy 113 NLRQOPKQPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEOLKSKTLFS 172
Db 141 SL-QAVSNFTAAYQESKREVLVFLFDGSSSKWAVEFDNKNKFIKDVKKLNSSIKFA 199
Qy 173 LMQYSEEPRIHFTKFEQNNPNRSLIKPTQLLGRTHATGIRKVVRELFN-ITNGARK 231
Db 200 AVQFSTEIRVDFDNDYQNGSABEKLME-RHMKSLTNTYKAINVYLKNVLSVSGADP 258
Qy 232 NPKIILITDGEKFGDPLQYED--VIPADREGVIRYVIGVGDAPRSEKSRQELNTVAS 289
Db 259 NAQKALVIITD-----GDPDNDNDYINLCDEQNILRYIIGV-----KVDLTTLTQLAA 309
Qy 290 KPRDHVFOINNFEALKTIQNLREKI FAIEGTQTGSSSFEHEMSQEGFAAITSNGPL 349
Db 310 EPKLANFTYIQEYNGUKGLDNLQKKIYNIEGSEKAHGRDRQKELSQSGSVVYQESVI 369
Qy 350 LSTVGSYDWAGGVFLYT---SKEKSTFINMTRVSDMDNDAYLGAAAIIILNRVQSLVGL 406
Db 370 VGSVGSNDRGALYEWGSGSKFRQTEITDPAN---KDSYMGYSTVLGRHGVSLFLSG 426
Qy 407 APRYQHIGLVAMPFRONTGMWESNANVKGTGTGAYFGASICSVDVDSNGSTDLVLIGAPHY 466
Db 427 APRAEHTGLVTLFTKNESTWTVMRNINEQIGSYFGASLSLLDVSDDGSDFLVLVGAFL 486
Qy 467 YE-QTRG-QGVSVCLPLPRQARWQCDVLYGEQ--GQPMGRFGAALTVLGDVNGDKLTD 522
Db 487 YSQPRAEGRLYVYTL-----SEQYSQKTLQSTTGRFATSLASLKDNLGDLSD 535
Qy 523 VAIGAPGEEDNRGAVYLFHTSGSGISPSHS-ORIASGKLSRLQYFGOSLGGQDLTMD 581
Db 536 VAVGAPLE--NEGVTYIYLDGTHGINPEHAPORIPARVLPGLQQLQFGVSLSGQMDND 593


```

DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020045; AAC23662.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; int_alpha; 4.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match 19.9% Score 1167.5; DB 11; Length 1167;
Best Local Similarity 29.0%; Pred. NO. 2e-77;
Matches 358; Conservative 214; Mismatches 456; Indels 205; Gaps 43;

QY 1 ENLDTEA--MTFQENARFGQSVVLOGSRVVGAPQEIIVANQGS-----LYQCDY 52
DB 20 FNMVDVMAWVTAALQGPVAVLSLLHLDPSN-----NOTCLLVARRSNRNTAALYRCAI 74
QY 53 STGSCPTIRLOVPEAVNMSGLSLAATT--SPOLLAC-GPTVHQTCSNTYKGLCLF 109
DB 75 SI-SPELTAQ-PVEHICMPKRGYQVTLVGNHNGVLVCIQVQARKFSLNSELTGACSL 132
QY 110 FGSNLRQOPKPEALRG-----C-----PQE 131
DB 133 LTPNLDLQAQVFSDFLEGFLDPGAHVDSGDYCRSGKSTGBEKKSAARRRTVEBEDEED 192
QY 132 DSDIAFLVDSGSIIPHFRAKERISTVMQL--KSKTILFSLMOYSEEFRIHPTKEF 189
DB 193 GTEIAVLDSGSGSIEPSDFQAKNFISTMRNFYEKCEFCNFALVQYGAIVQTEFDLQES 252
QY 190 QNNPNRSLPIITOLLGRTHATGIRKVVRELFTNGARKNAPKILILITDGEKFGDP 249
DB 253 RDINASLAKVOSIVQKEVKTASAMQHVLDNIFIPSRGSKKALKVMVLTGDIFGDP 312
QY 250 LGYEDVPEADREGVIRVIGVDFAFRSEKSRQELNTVASKPRDHVFQINNFEALKTQ 309
DB 313 LNLTVINSKMGVVRGAIGVDGAFKNNVTRELKLIASDPKAEHTKVTNYSALDGLL 372
QY 310 NQLEKIFAIEGTQGTSSSEHEMSQEGFSNAITSNGP-LLSTVGSVDWAGGVFLY-TS 367
DB 373 SKLQORIIVHMEGT--VGDALQYLAQTGFSAQILDKQVLGTGVAFNWSGGALLYSTQ 429
QY 368 KEKSTFIWNT-RVDS-DMNDVGLVAAAIILNRVQSLVGLAPRYOHIGLVAMPFRONTGM 425
DB 430 NGRGCFNLQTKEDSRTVOYSLGYSVLAVLHKHAGISIVAGAPRHUKRGAVFELKRED- 488
QY 426 WESNA--NVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVEQTRGGQVSCPLPR 482
DB 489 -EEDAFVRRIEGEQSGVFGSVLCPVDIDMDGTDFTDLVAAFPFYHIRGEGRVYVQVPE 547
QY 483 QORARWQCDVLYGQGGQPGWRFGAALTVLGVDNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSGHPGHTNSRFMGFAAAVGDINODKFTDVAIGAPLEGFGAGDGSYGS 606
QY 537 VYLFHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDILTMDGLVDLTVGAQGHVLL 596

```

RESULT 14

```

O88341
ID O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin alpha E2 (fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.2978 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481b-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMWEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq 19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5862	99.8	1153	19 AAW65090	Human Beta-integrin
2	5862	99.8	1153	21 AAB07360	Human CD11b protei
3	5862	99.8	1153	23 ABG61469	Human Beta2 integr
4	5862	99.8	1153	23 AAU80252	Human integrin l a
5	5862	99.8	1153	23 AAO14428	Integrin Mac-1 alp
6	5852	99.6	1153	11 AAR04136	Alpha subunit of M
7	3460	58.9	1153	11 AAR07120	p150.95 alpha subu
8	3446	58.6	1163	19 AAW65091	Human Beta-integrin
9	3446	58.6	1163	21 AAB07361	Human CD11c protei

10	3446	58.6	1163	23 ABG61470	Human Beta2 integr
11	3444	58.6	1163	24 ABU07406	Protein differenti
12	3411	58.0	1161	16 AAR78166	Human beta-2 integ
13	3411	58.0	1161	18 AAW23049	Human beta 2 integ
14	3411	58.0	1161	19 AAW72825	Human alpha-d. Ho
15	3411	58.0	1161	19 AAW65089	Human Beta2-integr
16	3411	58.0	1161	19 AAW57491	Human Beta2 integr
17	3411	58.0	1161	20 AAW73342	Human alphad prote
18	3411	58.0	1161	21 AAB07359	Human alpha d clon
19	3411	58.0	1161	23 ABG61468	Human Beta2-integ
20	3395.5	57.8	1161	18 AAW23064	Human beta 2 integ
21	3395.5	57.8	1161	19 AAW72837	Human alpha-d deri
22	3395.5	57.8	1161	19 AAW65106	Human Beta-integr
23	3395.5	57.8	1161	20 AAW73343	Human alphad prote
24	3395.5	57.8	1161	21 AAB07376	Human alpha d prot
25	3395.5	57.8	1161	23 ABG61485	Human Beta2-integ
26	3234.5	55.0	1161	16 AAR78169	Rat alpha-d subuni
27	3232.5	55.0	1161	18 AAW23062	Rat beta 2 integr
28	3232.5	55.0	1161	19 AAW72824	Rat alpha-d #1. R
29	3232.5	55.0	1161	19 AAW60004	Rat alpha d polype
30	3232.5	55.0	1161	21 AAB07374	Rat alpha d protei
31	3232.5	55.0	1161	23 ABG61483	Rat Beta2-integrin
32	3225.5	54.9	1161	19 AAW65104	Rat beta-integrin
33	3225.5	54.9	1161	20 AAW73345	Rat alphad protein
34	3222	54.8	1161	18 AAW23061	Mouse beta 2 integ
35	3222	54.8	1161	19 AAW72836	Mouse alpha-d #2.
36	3222	54.8	1161	19 AAW65103	Mouse beta-integr
37	3222	54.8	1161	19 AAW60003	Mouse alpha d poly
38	3222	54.8	1161	20 AAW73347	Mouse alphad prote
39	3222	54.8	1161	21 AAB07373	Mouse alpha-d prot
40	3222	54.8	1161	23 ABG61482	Mouse Beta2 integr
41	3218	54.8	1161	16 AAR78168	Mouse alpha-d subu
42	3211.5	54.7	1151	18 AAW23059	Rat beta 2 integr
43	3211.5	54.7	1151	19 AAW72834	Rat alpha-d #2. R
44	3211.5	54.7	1151	19 AAW65101	Rat beta-integrin
45	3211.5	54.7	1151	19 AAW60001	Rat alpha d polype

ALIGNMENTS

RESULT 1
AAW65090

ID AAW65090 standard; Protein; 1153 AA.

AC AAW65090;

DT 28-SEP-1998 (first entry)

DE Human Beta-integrin CD11b subunit protein.

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
lung inflammation; acute respiratory distress syndrome; CD11b subunit;
rheumatoid arthritis.

OS Homo sapiens.

PN US728533-A.

PD 17-MAR-1998.

PF 07-JUN-1995; 95US-0485618.

PR 07-JUN-1995; 95US-0485618.

PR 23-DEC-1993; 93US-0173497.

PR 05-AUG-1994; 94US-0286889.

PR 21-DEC-1994; 94US-0362652.

PA (ICOS-) ICOS CORP.

PI Gallatin WM, Van DER VIEREN M;

DR WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using

PT immobilised or labelled alpha-d polypeptide, useful for, e.g.

PT treating type-I diabetes

XX Example 5; Fig 1A-D; 106pp; English.

CC This sequence represents a human beta-integrin CD11b subunit which is

CC used to describe a method for identifying compounds that modulate the

CC interaction of the beta-integrin alpha-d subunit with a binding partner

CC of alpha-d which involves contacting an alpha-d polypeptide with an

CC alpha-d binding partner, one of which is immobilised and the other of

CC which is labelled, in the presence of a test compound, and determining if

CC the compound affects binding between the alpha-d polypeptide and alpha-d

CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment

CC comprising the cytoplasmic, transmembrane or extracellular domain of

CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,

CC asthma, psoriasis, lung inflammation, acute respiratory distress

CC syndrome and rheumatoid arthritis.

XX Sequence 1153 AA;

QY Query Match 99.8%; Score 5862; DB 19; Length 1153;

DB Best Local Similarity 99.3%; Pred. No. 0;

Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLOGSRVVVGAPOEIVAAQNRGLYQCDYSTGCEPI 60

DB 17 FNLDTENAMTFOENARFGQSVVQLOGSRVVVGAPOEIVAAQNRGLYQCDYSTGCEPI 76

QY 61 RLQVPVEAVNMLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLQPOK 120

DB 77 RLQVPVEAVNMLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLQPOK 136

QY 121 FPEALRGCPQESDIAFLVDGSGIIPHDFRAKEPISTVMEQLKSKTFLSLMQYSEF 180

DB 137 FPEALRGCPQESDIAFLVDGSGIIPHDFRAMEFVSTVMEQLKSKTFLSLMQYSEF 196

QY 181 RIHFTKPEQNNPNRSLKPIQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 240

DB 197 RIHFTKPEQNNPNRSLKPIQLLGRTHATGIRKVVRELFNITNGARKNAFKLIL 256

QY 241 TDGEKCPDLYEDVTPEADREGVIRYVGVGDAFSEKSRQELNTVASKPRDHVFQIN 300

DB 257 TDGEKCPDLYEDVTPEADREGVIRYVGVGDAFSEKSRQELNTVASKPRDHVFQIN 316

QY 301 NFPAKTIQNLREKIFAIEGTQSGSSFEHEMSOEGFSAATNSGPLLSTVGSYDWAG 360

DB 317 NFPAKTIQNLREKIFAIEGTQSGSSFEHEMSOEGFSAATNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILRNRVQSLVGLGAPRYQHIGLVAMFR 420

DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILRNRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQGVSCPL 480

DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQGVSCPL 496

QY 481 PRGORARWOCDAVLYGEQGPWGRFGAALTIVLDGVNGDKLTDVAIGAPGEEDNRGAVLYF 540

DB 497 PRGORARWOCDAVLYGEQGPWGRFGAALTIVLDGVNGDKLTDVAIGAPGEEDNRGAVLYF 556

QY 541 HGTSGSGISPSHSQIRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600

DB 557 HGTSGSGISPSHSQIRAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616

QY 601 PVLVRVAIMEFNPFEARNVPCNDQVVGKEAGEVRVCLHVQKSTRDRRLREGIQISVVT 660

DB 617 PVLVRVAIMEFNPFEARNVPCNDQVVGKEAGEVRVCLHVQKSTRDRRLREGIQISVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVRLNLF 720

DB 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVRLNLF 736

QY 721 SLVGTPLSAFQNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780

DB 737 SLVGTPLSAFQNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840

DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900

DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916

QY 901 QLELPVKYAVVMVTSHGCVSTKYLNFTASENTSRVMOHYOVSNLGRSLPISLVFLVPV 960

DB 917 QLELPVKYAVVMVTSHGCVSTKYLNFTASENTSRVMOHYOVSNLGRSLPISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020

DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036

QY 1021 PFGIQEENATLKGNSLSPDWIKYKSHNHLIVSTAEILFNDSVFTLLPGQGAFFRSQTET 1080

DB 1037 PFGIQEENATLKGNSLSPDWIKYKSHNHLIVSTAEILFNDSVFTLLPGQGAFFRSQTET 1096

QY 1081 KVEPEPVNPLPLIVGSSVGGLLALITALYKLGFEKROYKDMMSGGPPGAEPQ 1137

DB 1097 KVEPEPVNPLPLIVGSSVGGLLALITALYKLGFEKROYKDMMSGGPPGAEPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; Protein; 1153 AA.

XX

AC AAB07360;

XX

DT 17-JAN-2001 (first entry)

XX

DE Human CD11b protein sequence.

XX

KW Human; macrophage infiltration inhibition; alpha_d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11b.

XX

OS Homo sapiens.

XX

PN WO200029446-A1.

XX

PD 25-MAY-2000.

XX

PF 16-NOV-1999; 99WO-US27139.

XX

PR 16-NOV-1998; 98US-0193043.

PR 08-JUL-1999; 99US-0350259.

XX

PA (ICOS-) ICOS CORP.

XX

PI Gallatin MW, Van Der Vieren M;

XX

DR WPI; 2000-387751/33.

XX

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

PT macrophage infiltration and reduce inflammation at central nervous

PT system injury sites

XX

PS Example 5; Fig 1; 270pp; English.

XX

CC Integrins are a class of membrane-associated molecules that participate

CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11b. This
CC sequence was used in an alignment to identify a novel beta2 integrin
CC alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has
CC approximately 60% identity to the protein sequence of alpha d. The
CC Alpha d gene and protein may be useful in therapy for diseases linked
CC to aberrant alpha d function e.g. Type 1 diabetes, atherosclerosis,
CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
CC adhesion deficiency (LAD). In addition, anti-alpha d monoclonal
CC antibodies may be used in the inhibition of macrophage infiltration at
CC the site of a central nervous system injury. The monoclonal antibodies
CC can also be used to detect and diagnose Crohn's disease.
XX
SQ Sequence 1153 AA;
Query Match 99.8%; Score 5862; DB 21; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTQENARGFGQSVVOLQSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFGQSVVOLQSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATSPQQLACGTPVHTQTCSENTYVVKGLCFPGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPQQLACGTPVHTQTCSENTYVVKGLCFPGSNLRQPOK 136
Qy 121 FPEALRGCPQSDIAFLVDGSGSIIIPHDPRAKEFTSTVMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQSDIAFLVDGSGSIIIPHDPRAKEFTSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTFEFQNNPRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
Db 197 RIHFTFEFQNNPRSLVKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256
Qy 241 TDGSKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELNTVASKPPRHHVQIN 300
Db 257 TDGSKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELNTVASKPPRHHVQIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTCSSSFFHEMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTCSSSFFHEMSQEGFSAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWSNANVKGTOICAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVVCPL 480
Db 437 QNTGWSNANVKGTOICAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVVCPL 496
Qy 481 PRGQARWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREAVNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREAVNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVNFETKNSRRTQTQVLGLTQTCTETLKLQLENCIEDPVPVLRNLNF 720
Db 677 YDLALDSGRPHSRVNFETKNSRRTQTQVLGLTQTCTETLKLQLENCIEDPVPVLRNLNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG 780

Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNGNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGDSYETQVTFPPDLDSYRKVSTLQNRQSRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDGDSYETQVTFPPDLDSYRKVSTLQNRQSRQSRWLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPFENSESEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSESEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVVMVTSHGVTKYLNFTASNTSRVMQHQVQVSNLQSRSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVTKYLNFTASNTSRVMQHQVQVSNLQSRSLPISLVFLVPV 976
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1036
Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGOGAFVRSQTE 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGOGAFVRSQTE 1096
Qy 1081 KVBPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKXDMSEGPPGAEPO 1137
Db 1097 KVBPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKXDMSEGPPGAEPO 1153
RESULT 3
ABG61469
ID ABG61469 standard; Protein; 1153 AA.
XX ABG61469;
XX 27-AUG-2002 (first entry)
XX Human Beta2 integrin alphaCD11b subunit.
DE Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX Homo sapiens.
OS
XX WO200230980-A2.
PN
XX 18-APR-2002.
PD
XX 15-OCT-2001; 2001WO-US32059.
PF
XX 13-OCT-2000; 2000US-0688307.
PR
XX (ICOS-) ICOS CORP.
FA
XX Gallatin WM, Van Der Vieren M;
PI
XX MPI; 2002-463260/49.
DR
XX
XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -
XX
XX Example 5; Page 191-194; 270pp; English.
PS
XX The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an

anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a spinal cord injury victim. The method also involves the use of a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule, vascular cell adhesion molecule). The method is useful for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury. In particular, the spinal cord injury comprises compression of the spinal cord. The antibodies are also useful for reducing inflammation at the site of a central nervous system injury. The specification also details the identification of Beta2 integrin alpha2 CDNA and proteins, for use in raising the antibodies. Beta2 integrins are implicated in diseases such as LAD (leukocyte adhesion deficiency, inflammatory response, diabetes, multiple sclerosis, arthritis, graft atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, immune complex alveolitis and leukaemia. The present sequence is a Beta2 integrin alpha subunit sequence included for comparison with the Beta2 integrin alphaad protein sequences.

SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 23; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGQSVVQLQSGRVVVGAPQEIIVAAANQGRSLYQCDSYSGCEPI 60
 DB 17 FNLDTENAMTFQENARGQSVVQLQSGRVVVGAPQEIIVAAANQGRSLYQCDSYSGCEPI 76

QY 61 RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRLQRPQK 120
 DB 77 RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRLQRPQK 136

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKBFISTVMEQLKKSKTFLSLMYSEEF 180
 DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKBFISTVMEQLKKSKTFLSLMYSEEF 196

QY 181 RHFTFKFQNNPRLSLKPTTOLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
 DB 197 RHFTFKFQNNPRLSLKPTTOLLGRTHATGIRKVVRELFINITNGARKNAFKILIVI 256

QY 241 TGEKFGDPLGVEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTVASKPPDRHFQIN 300
 DB 257 TGEKFGDPLGVEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTVASKPPDRHFQIN 316

QY 301 NFEALXTIQNLRKIFAIEGTQSGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
 DB 317 NFEALXTIQNLRKIFAIEGTQSGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAILRNVRQSLVLCAPRYOHIGLVAMPR 420
 DB 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAILRNVRQSLVLCAPRYOHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEOTRGQGVSVCL 480
 DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEOTRGQGVSVCL 496

QY 481 PRQARWQCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
 DB 497 PRQARWQCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600
 DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 616

QY 601 PVLVRVKAIMEFNPVARNVFECDQVWKGKAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
 DB 617 PVLVRVKAIMEFNPVARNVFECDQVWKGKAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

QY 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDVPSPVLRNLF 720
 DB 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDVPSPVLRNLF 736

QY 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPFEKNCNDNICODDLSITFFSMCLDCLVWG 780
 DB 737 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPFEKNCNDNICODDLSITFFSMCLDCLVWG 796

QY 781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 DB 797 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
 DB 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916

QY 901 QLELPVKYAVYVMVTSHGVSSTKYLNFTASNTSRVMQHQYQVSNLQGRSLPISLVFLVPV 960
 DB 917 QLELPVKYAVYVMVTSHGVSSTKYLNFTASNTSRVMQHQYQVSNLQGRSLPISLVFLVPV 976

QY 961 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIP 1020
 DB 977 RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIP 1036

QY 1021 PFGIQEBSFNATLKNLSFDWYIKTSHNHLIIVTAEILFNDSVFTLLPGQGAFFVRSOTET 1080
 DB 1037 PFGIQEBSFNATLKNLSFDWYIKTSHNHLIIVTAEILFNDSVFTLLPGQGAFFVRSOTET 1096

QY 1081 KVEPFEPVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
 DB 1097 KVEPFEPVNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 4
 AAU80252
 ID AAU80252 standard; Protein; 1153 AA.
 AC AAU80252;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human integrin 1 alpha-M subunit protein.
 XX
 KW Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
 KW inflammatory disease; autoimmune disorder; Crohn's disease;
 KW human immunodeficiency virus; HIV; myocardial infarction;
 KW Sjorgen's syndrome; rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 499..500
 FT /note= "Encoded by GCG CAG AGG"
 FT
 XX WO200218583-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 31-AUG-2001; 2001WO-US27227.
 XX
 PR 01-SEP-2000; 2000US-229700P.
 XX
 PA (BLOO-) CENT BLOOD RES INC.
 XX
 PI Springer TA, Shimoaka M, Lu C;
 XX
 XX WPI; 2002-382964/41.
 DR N-PSDB; ABK50046.
 XX
 PT Modified integrin-I or integrin I-like domain polypeptide useful as an
 PT immunogen to produce antibodies specific to polypeptide, comprises a
 PT disulfide bond such that polypeptide is stabilized in a desired
 PT conformation
 XX
 PS Disclosure; Page 109-112; 112pp; English.
 XX
 CC This invention relates to a modified integrin-I or integrin I-like

CC domain polypeptide comprising at least one disulfide bond so that the
 CC domain is stabilised in a desired conformation. The polypeptide of
 CC the invention may have antiinflammatory or immunosuppressive activities.
 CC The polypeptides of the invention have an open conformation and are
 CC useful as immunogens to produce antibodies that selectively bind to
 CC integrin I-domain; and for identifying a modulator of integrin activity,
 CC or of interaction of an integrin and a cognate ligand. The polypeptide
 CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
 CC is useful for treating or preventing an integrin-mediated disorder which
 CC is an inflammatory or autoimmune disorder in a subject and for
 CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
 CC disease, nephritis; human immunodeficiency virus (HIV), myocardial
 CC infarction, Sjogren's syndrome, rheumatoid arthritis, dermatitis.
 CC A therapeutic composition comprising the peptide of the invention is
 CC useful for treating an integrin mediated disorder in a subject. The
 CC polypeptides and/or active or antigenic fragments are useful as
 CC reagents for diagnosis of integrin-mediated disorders. The present
 CC sequence represents the human integrin-1 alpha-M protein subunit used to
 CC generate the mutant polypeptides of the invention.
 XX
 SQ Sequence 1153 AA;

Query March 99.8%; Score 5862; DB 23; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDENAMTFQENARGFGQSVWLOGSRVVGVAPOEIVAAQNRGSLYQCDYSTGSCPEI 60
 Db 17 FNLDENAMTFQENARGFGQSVWLOGSRVVGVAPOEIVAAQNRGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 120
 Db 77 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFLFGSNLRQOPQK 136

Qy 121 PPEALRGCPQSDSIAFLVDGSGSIIIPHDPRAKEFISTVMEQLKSKTFLPSLMQYSEEF 180
 Db 137 PPEALRGCPQSDSIAFLVDGSGSIIIPHDPRAKEFISTVMEQLKSKTFLPSLMQYSEEF 196

Qy 181 RIHFTFEFQNNPNSRLIPITOLLGRTHATGIRKVVRELFNITGARKNAFKILLI 240
 Db 197 RIHFTFEFQNNPNSRLIPITOLLGRTHATGIRKVVRELFNITGARKNAFKILLI 256

Qy 241 TDGKFGDPLGYEDVIEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
 Db 257 TDGKFGDPLGYEDVIEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN 316

Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
 Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFTNMTRVDSMDNDVILGYAAAIIILNRVQSLVLAGAPRYOHIGLVAMER 420
 Db 377 GVFLYTSKEKSTFTNMTRVDSMDNDVILGYAAAIIILNRVQSLVLAGAPRYOHIGLVAMER 436

Qy 421 QNTGWSNANVKGTQIGAFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCPPL 480
 Db 437 QNTGWSNANVKGTQIGAFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCPPL 496

Qy 481 PRGQARWQCDAVLYGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
 Db 497 PRGQARWQCDAVLYGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYVLF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGGQDLTMDGLVDLTVGAGQHVILLRSQ 600
 Db 557 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGGQDLTMDGLVDLTVGAGQHVILLRSQ 616

Qy 601 PVLRVKAIMENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
 Db 617 PVLRVKAIMENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
 Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 736

Qy 721 SLVGTPLSAGNLRPVLAEADAQRLFTALPFPFKNCGNDNICODDLSITFSFMSLDCLVWG 780
 Db 737 SLVGTPLSAGNLRPVLAEADAQRLFTALPFPFKNCGNDNICODDLSITFSFMSLDCLVWG 796

Qy 781 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACSSASSTEV 840
 Db 797 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACSSASSTEV 856

Qy 841 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
 Db 857 SGALKSTSCSINHPIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 916

Qy 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSUPISLVFLVPV 960
 Db 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSUPISLVFLVPV 976

Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1020
 Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1036

Qy 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFFRSQTET 1080
 Db 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFFRSQTET 1096

Qy 1081 KVEPFFVNPPLPIVGVSSVGLLLALITAAALYKLGFFKROYKDMKSEGGPPGAEPQ 1137
 Db 1097 KVEPFFVNPPLPIVGVSSVGLLLALITAAALYKLGFFKROYKDMKSEGGPPGAEPQ 1153

RESULT 5
 AAO14428
 ID AAO14428 standard; protein; 1153 AA.
 XX AAO14428;
 AC AAO14428;
 XX
 DT 03-MAY-2002 (first entry)
 XX
 DE Integrin Mac-1 alpha subunit.
 XX
 KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
 KW open conformation; integrin related inflammatory disorder;
 KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
 KW reperfusion; hypovolemic shock; infarction; cerebral shock;
 KW viral infection; cancer; gene therapy; vaccine;
 KW bioactive agent screening.
 XX
 OS Unidentified.
 PN WO200204521-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 09-JUL-2001; 2001WO-US21805.
 XX
 PR 07-JUL-2000; 2000US-216600P.
 XX
 PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 PA (BLOO-) CENT BLOOD RES.
 XX
 PI Springer T;
 XX
 DR WPI; 2002-148167/19.
 XX
 PT New integrin I domain protein having alteration in at least 2
 PT noncontiguous regions and exits in an open conformation, useful for
 PT treating, preventing or suppressing inflammatory or immunological
 PT disorders
 XX
 PS Example 1; Fig 1F; 90pp; English.
 CC
 XX The invention comprises structurally biased variant integrin inserted (I)
 CC domain proteins, wherein the alterations to the protein occur in at least

CC two noncontinuous regions. Specifically the the variant integrin I domain
 CC proteins are structurally biased to exist in the open conformation.
 CC thereby altering the binding ability of the protein. The invention also
 CC comprises nucleic acids encoding the variant integrin I domain proteins.
 CC The integrin I domain proteins and nucleic acids are useful for treating,
 CC preventing or suppressing integrin related inflammatory and immunological
 CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain
 CC proteins and nucleic acids can also be used for treating: ischaemia/
 CC reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
 CC infection; and cancer. The variant integrin I domain nucleic acids and
 CC proteins may be used in gene therapy, as vaccines and to screen for
 CC bioactive agents. The present amino acid sequence represents the Mac-1
 CC alpha subunit of integrin.

SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 23; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFOENARGFGOSVVQLOGSRVVVGAQOEIVAAQNRGLYQCDYSTGSCBPI 60
 DB 17 ENLDTENAMTFOENARGFGOSVVQLOGSRVVVGAQOEIVAAQNRGLYQCDYSTGSCBPI 76
 QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVVGKLCFLFGSLNRQQPOK 120
 DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVVGKLCFLFGSLNRQQPOK 136
 QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFAKRAKEPISVMEQLKKSKTLFLSLMQYSEEF 180
 DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFAKRAKEPISVMEQLKKSKTLFLSLMQYSEEF 196
 QY 181 RHFTPEKFNQNPRLSKPTOLLGRTHATGIRKVVRELNFNITNGARKNAFKLIL 240
 DB 197 RHFTPEKFNQNPRLSKPTOLLGRTHATGIRKVVRELNFNITNGARKNAFKLIL 256
 QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVGVGDPAFSEKSRQELNTVASKPRDRHVFQIN 300
 DB 257 TDGEKFGDPLGVEDVPEADREGVIRYVGVGDPAFSEKSRQELNTVASKPRDRHVFQIN 316
 QY 301 NFALKTIQNLREKIFAIEGTOTGSSSFHEHMSQEGFSAITNSGPLLSTVGSYDWAQ 360
 DB 317 NFALKTIQNLREKIFAIEGTOTGSSSFHEHMSQEGFSAITNSGPLLSTVGSYDWAQ 376
 QY 361 GVFLYTSKEKSTFNTRVDSMDNDAYLGVAIAIILNRVOSIVLGAPYOHIGLVAMFR 420
 DB 377 GVFLYTSKEKSTFNTRVDSMDNDAYLGVAIAIILNRVOSIVLGAPYOHIGLVAMFR 436
 QY 421 QNTGMHESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 480
 DB 437 QNTGMHESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 496
 QY 481 PRQARWOCDAVLYGEQGPWRGFAALTVLGDVNGDKLTDVAIGAPGEENRGAIVLYF 540
 DB 497 PRQARWOCDAVLYGEQGPWRGFAALTVLGDVNGDKLTDVAIGAPGEENRGAIVLYF 556
 QY 541 HGTSGSGISPHSQRTAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVCAGHVLILRSQ 600
 DB 557 HGTSGSGISPHSQRTAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVCAGHVLILRSQ 616
 QY 601 PVLVRKAIMFENPREVARNVFCNDQVVGKEAGEVRVCLHVQKSTRDRLRREQIQSVVT 660
 DB 617 PVLVRKAIMFENPREVARNVFCNDQVVGKEAGEVRVCLHVQKSTRDRLRREQIQSVVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSTRRTQVLGLTQTCTETLKLQLPNCIEDPVSPIVRLNLF 720
 DB 677 YDLALDSGRPHSRVAFNETKNSTRRTQVLGLTQTCTETLKLQLPNCIEDPVSPIVRLNLF 736
 QY 721 SLVGTPLSARGNLRPVLAEDAORLFTALPFPEKNCNDNITCODDLSTITFSFMSLDCLVWG 780
 DB 737 SLVGTPLSARGNLRPVLAEDAORLFTALPFPEKNCNDNITCODDLSTITFSFMSLDCLVWG 796
 QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840

DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
 DB 857 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
 QY 901 QLELPVKYAVYVMVYVTSKYNLFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 960
 DB 917 QLELPVKYAVYVMVYVTSKYNLFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 976
 QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1020
 DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1036
 QY 1021 PFGIQEEFNATLKGNLSPFDWYIKTSHNHLITVSTAEILFNDSVFTLLPGQGAFVRSOTET 1080
 DB 1037 PFGIQEEFNATLKGNLSPFDWYIKTSHNHLITVSTAEILFNDSVFTLLPGQGAFVRSOTET 1096
 QY 1081 KVEPFEVPNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKQDMSEGPPGAEPPQ 1137
 DB 1097 KVEPFEVPNPPLIVGSSVGGLLLALITAAALYKLGFFKQYKQDMSEGPPGAEPPQ 1153

RESULT 6
 AAR04136
 ID AAR04136 standard; protein; 1153 AA.
 XX AAR04136;
 XX 25-MAR-2003 (updated)
 DT 07-SEP-1990 (first entry)
 XX Alpha subunit of Mac-1 leukocyte adhesion receptor.
 XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
 KW non-specific defence system; integrin gene superfamily.
 XX synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 86..88
 FT /*label= putative N-glycosylation site
 FT Modified-site 240..242
 FT /*label= putative N-glycosylation site
 FT Modified-site 391..393
 FT /*label= putative N-glycosylation site
 FT Modified-site 469..471
 FT /*label= putative N-glycosylation site
 FT Modified-site 693..695
 FT /*label= putative N-glycosylation site
 FT Modified-site 697..699
 FT /*label= putative N-glycosylation site
 FT Modified-site 735..737
 FT /*label= putative N-glycosylation site
 FT Modified-site 802..804
 FT /*label= putative N-glycosylation site
 FT Modified-site 881..883
 FT /*label= putative N-glycosylation site
 FT Modified-site 901..903
 FT /*label= putative N-glycosylation site
 FT Modified-site 912..914
 FT /*label= putative N-glycosylation site
 FT Modified-site 941..943
 FT /*label= putative N-glycosylation site
 FT Modified-site 947..949
 FT /*label= putative N-glycosylation site
 FT Modified-site 979..981
 FT /*label= putative N-glycosylation site
 FT Modified-site 994..996
 FT /*label= putative N-glycosylation site
 FT Modified-site 1022..1024
 FT /*label= putative N-glycosylation site

OS Homo sapiens.
 PN WO200029446-A1.
 XX 25-MAY-2000.
 XX 16-NOV-1999; 99WO-US271139.
 XX 16-NOV-1998; 98US-0193043.
 PR 08-JUL-1999; 99US-0350259.
 XX (ICOS-) ICOS CORP.
 PA Gallatin MW, Van Der Vieren M;
 PI WPI; 2000-387751/33.
 XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
 PT macrophage infiltration and reduce inflammation at central nervous
 PT system injury sites
 XX Example 5; Fig 1; 270pp; English.
 XX Integrins are a class of membrane-associated molecules that participate
 CC in cellular adhesion. Integrins are made up of an alpha subunit and a
 CC beta subunit. One class of human integrins are restricted to expression
 CC in white blood cells and have a common beta2 subunit: the leukocyte
 CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
 CC have an important role in immune and inflammatory responses. The present
 CC protein sequence is the human integrin alpha subunit CD11c. This
 CC sequence was used in an alignment to identify a novel beta2 integrin
 CC alpha subunit: alpha_d (AA66014 and AAB07359). The present sequence has
 CC approximately 66% identity to the protein sequence of alpha_d. The
 CC Alpha_d gene and protein may be useful in therapy for diseases linked
 CC to aberrant alpha_d function e.g. Type I diabetes, atherosclerosis,
 CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
 CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
 CC adhesion deficiency (LAD). In addition, anti-alpha_d monoclonal
 CC antibodies may be used in the inhibition of macrophage infiltration at
 CC the site of a central nervous system injury. The monoclonal antibodies
 CC can also be used to detect and diagnose Crohn's disease.
 XX Sequence 1163 AA;
 SQ Query Match 58.6%; Score 3446; DB 21; Length 1163;
 Best Local Similarity 60.8%; Pred. No. 1e-278;
 Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;
 QY 1 FNLDTENAMTFQENARGFGOSVVOLOQSRVVVGAPQEIIVANORGSLVQCDYSTGSGCEPI 60
 DB 20 FNLDTEELTAFRVDSAGFSDSVVOYANSVWVGAPQKIIAANOIGGLYQCYSTGACEPI 79
 QY 61 RLQVPVAVNMSLGLSLAATTPSPQLLACGPTVHQTCSNTYVYKGLCFGLFSGNLROQPOK 120
 DB 80 GLQVPPVAVNMSLGLSLAATTPSPQLLACGPTVHHEGCRNMYLTGLCFLLGPT-QLTQR 137
 QY 121 FPEALRCGPQEDSDIAELVDSGSIIPHDPRRAKEPISTVNEQLKKSKTLFSLMQSEEP 180
 DB 138 LPVSRQECPRQEQDIFVLIDGSGSISRNFATMNFRAVISQFQRPSTQFSLMQSNKF 197
 QY 181 RIHETFEQNNPNRSLIPITQLLGRTHATGIRKVVRELENTINGARKNAFKILILI 240
 DB 198 QTHETFEFRRTSNPLSLASVHQLOQFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257
 QY 241 TDGKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQELNVTASPPRDHVFQIN 300
 DB 258 TDGKKGSDLYDKDVIHMAAAGIIRVIGVGLAFQFNNSWKELNDIASKPSQEHIFKVE 317
 QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHENSQEGFSAATTSNGLPLSTVGSYDWAG 360
 DB 318 DFDALKDIONQLREKIFAIEGTETISSSSFELEMAQEGFSAVTPDGPVLGAVGVSFTWSG 377
 QY 361 GVFLYTSKESKSTFNMTRVDSMDNDAYLVGAAILLNRVQSLVGLGAPRQHIGLVAMFR 420

DB 378 GAFLYPPNNSPTFINMSQENMDRSDYLGSTYELALWKGVSQSLVGLGAPRQHIGKAVIFI 437
 QY 421 QNTGMWESNANVKGTOICAYFGASLCSVDVDSNGSTDLVLCAPHYEQTGGQVSVCP 480
 DB 438 QVSRQWRMKAIEVIGTIGSYFGASLCSVDVDTGSDTLVLCAPHYEQTGGQVSVCP 497
 QY 481 PRGQARWQCCDAVLYGEOQPGWREGAALTVLGVDVNGDKLTDVAIGAPEEDNRGAVLYF 540
 DB 498 PRGWR-RWVCDVLYGEOGHPWREGAALTVLGVDVNGDKLTDVVGAPGEBENRGAVLYF 556
 QY 541 HGTSGSIGSPHSQRIAGSKLSPLRQYFGQSLGSGQDLTMDGLVDLTVGAAQHVLRLRSQ 600
 DB 557 HCVLGFSPHSQRIAGSKLSPLRQYFGQSLGSGQDLTMDGLVDLAVGARGQVLLLRTR 616
 QY 601 PVLRYKATMEFNPREVARNVPECNDVNVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
 DB 617 PVLVGVSNQFIPAEIPRSAFECEQVVSSEQTLVQSNICLYIDKSKNLLGSRDQSSVT 676
 QY 661 VDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
 DB 677 LDLALAPGLSPRAIFQETKNSLSRVVLGLKHCENFNLLPSCVEDSVPIILRLNF 736
 QY 721 SLVGTPLSAFGNLRPVLAEDAQRFTALFPPEKNCNDNI CODDLSITFSMSLDCLVVG 780
 DB 737 TLVKGPLLAFLNLRPMLAALAQRFTASLPPEKNCADHICQDNLGIGSFPGLKSLVVG 796
 QY 781 GPREFNVTVVNDGEDSVRTQVTFPFLDLSYRKVSTLQNRORSORSWRLACESASSTEV 840
 DB 797 SNLELNAEVMVWMDGEDSYGTTITFSHPAGLSRYVAEGQKQGLRSLHLTC--CSAPVG 854
 QY 841 SCALKSTCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
 DB 855 SGTWSTSCRIHNLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRTSKTIF 914
 QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFATAS-ENTSRVMQHYQVSNLQKSLPISLVFLVP 959
 DB 915 QLELPVKYAVYIVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
 QY 960 VRLNQTIVDRQVTFSENLSTCHTKERLPSHDFLAELKAPVNVCSINVCQICQDI 1019
 DB 975 VELNQEAVMVDEVSHPQNPQLRCSSEKIAPPASDFLAHQKPNVLDSCSIAGCLFRCDV 1034
 QY 1020 PFGLOEENATLKGNLSPDWIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSQTE 1079
 DB 1035 PSFSVOEELDTLKGNLSEFWVRQILQKKVSVVSAEILFDTSVTSQLPQCAFRAQTI 1094
 QY 1080 TKVEPPEVPNPLPIVGVSGGLLLALITAAALYKLGFFPKQYKQMMSE 1128
 DB 1095 TVLEKYKHNPILIVGSSIGLLALLITAVILYKVGFFPKQYKEMWEE 1143
 RESULT 10
 ABG61470
 ID ABG61470 standard; Protein; 1163 AA.
 XX AC ABG61470;
 XX DT 27-AUG-2002 (first entry)
 XX DE Human Beta2 integrin alphaCD11c subunit.
 XX KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
 KW LAD; leukocyte adhesion deficiency; inflammation response; diabetes;
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
 KW locomotor recovery; locomotor damage; locomotor impairment;
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
 OS Homo sapiens.

PF 08-APR-2002; 2002WO-US10824.
XX
PR 06-APR-2001; 2001US-281731P.
PR 06-APR-2001; 2001US-281732P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Jay G;
XX
DR WPI; 2003-058520/05.
XX
XX Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
PT patients -
PS
PS Claim 1; Page 225-228; 416pp; English.
XX
XX The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.
XX
SQ Sequence 1163 AA;

Query Match 58.6%; Score 3444; DB 24; Length 1163;
Best Local Similarity 60.7%; Pred. NO. 1.5e-278;
Matches 685; Conservative 139; Mismatches 299; Indels 6; Gaps 4;

QY 1 FNLDTEANMTFQBNARGFGQSVVQLOGSRVVVGAPOEIVAAANORGSLYCDYSTGSCPEI 60
DB 20 FNLDTEELTAFRVDSAGFQSVVQYANSVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTPSPQLLAGCTVHTQCSNTYVYKGLCFGLGNSLRQPOK 120
DB 80 GLQVPVEAVNMSLGLSLAATTPSPQLLAGCTVHTQCSNTYVYKGLCFGLGNSLRQPOK 137
QY 121 FPAALGCCPQEDSDIAFLVDGSGSIIPHDPRAKEFISTWMEQLKSKTLFSLMOYSEEF 180
DB 138 LPVSRQCEPRQEDQIVFLIDGSGSIISRNFAFMNFRVAVISQFQPSQFSLMQFSNKP 197
QY 181 RIHFTKFEQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240

DB 198 QTHLTFFEEFRRTSNPLSLASVHQLQOFTYTATAIQNVVHRLPHASYGARRDATKILIVI 257
QY 241 TDGEKFGDPLGYEDVIEADREGVIRYVIGVGDAFRSEKSRQBELNTVASKPRDHVQIN 300
DB 258 TDGKEGDTLDYKDVIPMADAAGIIRYAGVGLAFQNRNSWKELNDIASKPSQEHFKVE 317
QY 301 NFPAKTIQNLREKIFAIEGTOTGSSSPFEHMSQEGFSAALTSNGPLLSSTVGSVDWAG 360
DB 318 DFDALDIOQLREKIFPIEGTETSSSFELEMAQEGFSAVTFDGPVLGAYGSGFTMSG 377
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLVGAAGIILNRVQSLVGLAPRYQHLGLVAMFR 420
DB 378 GAFLYPPNMGPTFINMSQENVDNRDSYLGSTELALWGVQSLVGLAPRYQHTGKAVIFT 437
QY 421 QNTGMWESNANVKGITQIGFYGASLCSVDVDSNGSTDVLVIGAPHYYEQTQGGQSVCPPL 480
DB 438 QVSRQWRMKAETGTQIGSYFGPSLCSVDVDSGSTDVLVIGAPHYYEQTQGAQSVCPPL 497
QY 481 PRGORARQCDVLYGEGQGPWRFGAALTLDVGVNGDKLTDVAIGAPGEDNRGAYLFL 540
DB 498 PRGWR-RWVCDVLYGEGQHPWRFGAALTLDVGVNGDKLTDVVIGAPGEENRGAYLFL 556
QY 541 HGTSGSGISPSHSORIASKLSPLQYFGOSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGVLGSPISPSHSORIASQSLSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616
QY 601 PVLKVAIMEFNPREVARNVFECDQVQVKGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLWGVGVMQFPAEIPRSAFECEQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDGRPHSRVAVNETKNSRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 677 LDALDPLGRLSPRATFQETKNSRVRVLGLKAKCENFNLLPSCVEDSVTPITRLNF 736
QY 721 SLVGTPLSAFQNLPLAEDAQRLFTALFPEKCKGNDNICODDLSITFSPMSLDCLVVG 780
DB 737 TLVGKPLLAFLNRLPLAADAQRYFTASLPEKCKGADHICQDNLGIFSFPGLKSLLVG 796
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSRWLACESASSTEV 840
DB 797 SNLEUNAEVWVNDGEDSYGTITFSPHAGLSYRYVAEGQKQGLRSLHLTCDAPVG- 854
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
DB 855 SQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVLGDRLLLTANVSSENTPRTSKTF 914
QY 901 QLELPVKYAVYVTVSHCVSTKYLNFAS-ENTSRVMOHOYOVSNLQORSLSILVFLVP 959
DB 915 QLELPVKYAVYVTVSSHQEQFTKYLNFSESEKESHVAMHRYOVNMLGQDLPVSNFWVP 974
QY 960 VRLNQTVIDRPOVTFSNLSSTCHTKERLPSHSDFLAELRKAAPVYVNCISVACQICQDI 1019
DB 975 VELNQEAVMQVDEVSPLQNPSPSRCSSEKIAGPSDFLAHIQKNPVLDCSIAGCLFRCDV 1034
QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSTLPGQAFVRSQTE 1079
DB 1035 PSFSVQEEELDTLKNLSFGVVRQILQKKVSVVAEITFTDTSVYSQLPQGEAFRAQTT 1094
QY 1080 TKVEPFEVNPPLPLIVGSSVGLLALLIITAAALYKLGFFKQYKQDMSE 1128
DB 1095 TVLEKYKVHNPETPLIVGSSIGLLALLIITAVLYKVGFQKQYKEMMEE 1143

RESULT 12
AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX
AC AAR78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX

XX KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
 KW immunoglobulin; fusion protein; binding molecule; antibody;
 KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
 KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
 KW acute respiratory distress syndrome.
 XX OS Homo sapiens.
 XX US5831029-A.
 XX PD 03-NOV-1998.
 XX PF 07-JUN-1995; 95US-0482293.
 XX PR 07-JUN-1995; 95US-0482293.
 XX PR 23-DEC-1993; 93US-0173497.
 XX PR 05-AUG-1994; 94US-0286889.
 XX PR 21-DEC-1994; 94US-0362652.
 XX PA (ICOS-) ICOS CORP.
 XX PI Gallatin WM, Van Der Vieren M;
 XX WI; 1998-609318/51.
 XX DR N-PSDB; AAV67281.
 XX PT Antibodies specific for beta-2 integrin alpha-subunit d - useful in
 XX immunohistochemical analysis
 XX Example 5; Column 61-66; 106pp; English.
 XX The present sequence represents human alpha d. The present invention
 CC describes: (1) an antibody that specifically binds alpha d, which
 CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
 CC antibody as in (1); (3) an anti-idiotypic antibody specific for the
 CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
 CC monoclonal antibody of (2). Antibodies specific for alpha d can be
 CC used in immunohistochemical analysis to localise alpha d to subcellular
 CC compartments or individual cells within tissues. Substances that modulate
 CC alpha d binding (which may include antibodies) can be used to treat
 CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
 CC inflammation, acute respiratory distress syndrome or arthritis.
 XX SQ Sequence 1161 AA;
 Query Match 58.0%; Score 3411; DB 19; Length 1161;
 Best Local Similarity 59.4%; Pred. No. 8.9e-276;
 Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;
 QY 1 ENLDTENAMTFQENARGFGOSVVOLOGSRVVGAPQEIIVAAVNOGSLYOCDSYSGCEPI 60
 DB 17 FNLDTBEPTIFQSDAGFGOSVVOFGSRVLVVGAPLEVAANOTGRLYDCAATGMCQPI 76
 QY 61 RLQVPVAVNMSLGLSAAATSPPLQACGFTVHOTCSENTYVYKGLCFGLSGNLROQPOK 120
 DB 77 PLHIREAVNMSLGLTAASTNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-IIIQT 135
 QY 121 FPALRGCPQEDSDIAPLVGSGSIIPHDPRRAKEFISTVNEQLKXSKTLFSLMQYSEEF 180
 DB 136 VPDATPECPHQEMDIIVFLDGGSIDQNDPNQMGFVQAVMGQFEGDTLFLALMQYSNLL 195
 QY 181 RIHTPEKFNQNNPSRLIKPIITQLLGRTHATGIRKVVRELENTINGARKNAFKILILI 240
 DB 196 KIHFTTFQRTFSQOSLVDPIVOLKGLTFTATGILTVTQLFHHKNGARKSAKKILIVI 255
 QY 241 TDCEKFDPLGYEDVPEADREGVIRYVIGVGAFAFRSEKRSQBELNTVASKPPRDHVFQIN 300
 DB 256 TDGQKYKDPLEYSDVIPAQKAGIIRYALGVGHAFQPTARQELNIISSAPPDHFVKVD 315
 QY 301 NFPAKTIQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
 DB 316 NFAALGSIQOLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMQGLFLGAVGSFSWSG 375

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 420
 DB 376 GAFLYPPNMSPTFINNSQENVDNRDSYLGSTELALWKGVQNLVLCAPRYQHTKAVIFT 435
 QY 421 QNTGMWESNANVKGTOIGAYFCASICSVDVDSNGSTDLVLCGAPHYETRGQVSVCP 480
 DB 436 QVSRQWRKKAETGTQIGSYFGASICSVDVDSNGSTDLVLCGAPHYETRGQVSVCP 495
 QY 481 PRQORARWQCDVAVLYGEOQPMGRFGAALTVLVDVNGDKLTDVAICAPGEENRGAVYLF 540
 DB 496 PRQORVQWQCDVAVLRGEQGHMGRFGAALTVLVDVNGDKLTDVAICAPGEENRGAVYLF 555
 QY 541 HGTSGSISPSHSQRIAGSKLSPRIQYFQSGISGQDLTMDGLVDLTVGAQGHVLLRSQ 600
 DB 556 HGASESGISPSHSQRIAGSKLSPRIQYFQSGISGQDLTMDGLVDLTVGAQGHVLLRSQ 615
 QY 601 PVLRYKATMEFNPVARNVFCNDQVVKGEAGEVVRVCLHVOKSTRDRLRQGLSVVT 660
 DB 616 PVLKVGAMRFPSPVEKAVYRCWBEKPSALEAGDATVCLTIQKSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRVFNENETKSTRRTQTVLGLTQCTETLKLQLPNCIEDPVPVILRLNF 720
 DB 674 FDLALDPGLTSLRAIFNETKNPTLTKRGLGIHCETLKLPLDCVEDVSPVILHLNF 733
 QY 721 SLVGTPLSAFGNLRPVLAEQAORLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCVLVVG 780
 DB 734 SLVREPIPSQNLRPVLAVGSQDLFTASLPFPEKNCQDGLCEGLDGLVTLFSGQLTLTVG 793
 QY 781 GPREFNVTYVNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSWLACESASTEV 840
 DB 794 SSLELNVIYVWVNGEDSYGTVSYLYPAGLSHRRVSGAKQPHOSALACETV-PTED 852
 QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSSNNMPRTNKTEF 900
 DB 853 EG-LSSRCVSNVHPFHEGSGNCTFVTFDVSYKATLGDRLMLRASSENKASSSKATF 911
 QY 901 QLELPKYAVYVMVTSVSHGVSTKYLNF-TASENTRVMQHOYQVSNLQSRSLPISLVLPV 959
 DB 912 QLELPKYAVYTMISRQEESTKYFNATSDKKKEAEHRYRVNLSQDLAISINFWVP 971
 QY 960 VRLNQTVIWDROVTFSENLSSTCHTKERLPSHDSFLAELRKAPVNVCSIAVQCORIQCDI 1019
 DB 972 VLLNGVAVWVYMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQPCRDV 1029
 QY 1020 PFEGIQEEFNATLKGNLSPDWIKTSHNHLIVSTAEILFNDSVFTLLPQCGAFVRSQTE 1079
 DB 1030 PFSVQOEELDTLKGNLSEGWVRETLQKVLVSVVAEITFTSVYSQLPQCGAFVRSQTE 1089
 QY 1080 TKVEPPEVNPPLIVGSSVGGLLIALITALYKLGFFKRYQKDMWSE 1128
 DB 1090 MVLDEDEVYNAIPIINGSSVGALLLALITATLYKLGFFKRYKEMLED 1138
 RESULT 15
 ID AAW65089
 XX AAW65089 standard; Protein; 1161 AA.
 AC AAW65089;
 XX 28-SEP-1998 (first entry)
 XX Human Beta-integrin alpha-d subunit protein isolated from clone 19A2.
 XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome;
 XX rheumatoid arthritis.
 OS Homo sapiens.
 XX Key
 XX Peptide 1..16
 FT

```

FT Protein /label= signal
FT 17..1161
FT /note= "Beta-integrin alpha-D subunit"
FT Domain 17..1108
FT /label= extracellular
FT Domain 150..352
FT /note= "homologous to insertion domain of CD11 subunits"
FT Domain 1129..1161
FT /label= cytoplasmic
XX
XX US5728533-A.
XX
XX 17-MAR-1998.
XX
XX 07-JUN-1995; 95US-0485618.
XX
XX 07-JUN-1995; 95US-0485618.
XX 23-DEC-1993; 93US-0173497.
XX 05-AUG-1994; 94US-0286889.
XX 21-DEC-1994; 94US-0362652.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gallatin WM, Van DER VIEREN M;
XX
XX WPI; 1998-206565/18.
XX N-PSDB; AAV35236.
XX
XX Screening assay for modulators of integrin binding - using
XX immobilised or labelled alpha-d polypeptide, useful for, e.g.
XX treating type-I diabetes
XX
XX Example 5; Fig 1A-D; 106pp; English.
XX
XX This sequence represents a novel human beta-integrin alpha-d subunit.
XX This sequence is used in a method for identifying compounds that modulate
XX the interaction of alpha-d with a binding partner of alpha-d which
XX involves contacting an alpha-d polypeptide with an alpha-d binding
XX partner, one of which is immobilised and the other of which is labelled,
XX in the presence of a test compound, and determining if the compound
XX affects binding between the alpha-d polypeptide and alpha-d binding
XX partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat
XX diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
XX asthma, psoriasis, lung inflammation, acute respiratory distress
XX syndrome and rheumatoid arthritis.
XX
XX Sequence 1161 AA;
XX
XX Query Match 58.0%; Score 3411; DB 19; Length 1161;
XX Best Local Similarity 59.4%; Pred. No. 8.9e-276;
XX Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;
XX
XX 1 FNLDTENAMTFQENARGFQGVVQLGGSRVVVGAPOEIVAAHQSLYQCDYSTGSCPEI 60
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 17 FNLDVEEPTIFEDAGGFGQGVVQGGSLVVGABLEVVAANQTLRVDCAATGWCQPI 76
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX 61 RLQVPVEAVNMSGLSLAATSPQLACGPTVHQTCSNTYVKGCLFLGSLNRQQPQK 120
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX 77 PLHIRPEAVNMSGLTLAASNGSRLLACGPTLHRVCGENSYSKGSCLLLGSRW-ELIQT 135
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX 121 FPEARLGGCPQEDSIAFLVDGSGSIIPHDPRAKEFISTVMEQLKKSKTLFLSMQYSEF 180
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX 136 VPDATPECPHQEMDIVFLFDGSGSIDQDNDFNMKGFFVQAVMGQFEGTDTLFLALMQYSNLL 195
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX 181 RHFTFKFQNNPRSLIKPTOLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX 196 KIHFTFTQRTSPQSLVDPIVOLKGLTFTATGILTVTQFLPHHNGARKSAKILIVI 255
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX 241 TDGEKFGDPLGVEDYIPEADREGVIRYIVGVGDAFRSEKSRQELNTVASKPPRDHVQIN 300
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX
XX 256 TDGQYKDPLEYSVDVIPAQEKAGIIRYAIYGVGHATQGPARTQELNTISSAPPQDHVKVD 315
XX |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY 301 NPEALKTIONQLREKIFAIEGTOTGSSSSPEHEMSQEGFAAITSNGPLLLSTVGSYDWAG 360
QY 316 NPAALGSIQLOLEKIYAVEGTQSRASSSFQHEMSQEGFTALTMDGLFAGVGSFWSG 375
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 420
QY 376 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSTELALWKGVQNLVLGAPRYOHTGKAVIFT 435
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGGGVSCPL 480
QY 436 QVSRQWRKKAIEVTGTQIGSYFGASLCSVDVDSNGSTDLILIGAPHYYEQTGGGVSCPL 495
QY 481 PRGORARWQCDAVLYGEOGQPMGRFGAALTVLGVDVNGDKLTDVAIGAPEGEDNRGAYVLF 540
QY 496 PRGORVQWQCDAVLYGEOGQPMGRFGAALTVLGVDVNGDKLTDVAIGAPEGEDNRGAYVLF 555
QY 541 HGTSGSISPSHSQRIAGSKLSPLRYQFGQSLSGQDLTMDGLVDLTVGAGHVLRLRSQ 600
QY 556 HGASEGISPSHSQRIASSQLSPRLQYFGQALSGQDLTQDGLMDLAVGARGVLLRLSL 615
QY 601 PVLRYKAIMFENPREVARNVPECNDQVVKGEAGEVRVCLHVOKSTDRLRREGIOISVVT 660
QY 616 PVLKVGVMRFSPVEVAVKVRCEWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSVR 673
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQPNCIEDPVSPVILRLNF 720
QY 674 FDLALDPGRITSRRAIFNETKNPTLTRRKTGLGHCETLKLKLLPDCVEDVSPILHLNF 733
QY 721 SLVGTPLSAFNGLRPLVAEDAQRFTALFPPEKNCNDNTICQDLSITFSFMSLDCLVWG 780
QY 734 SLVREPIPSQNLRLPVLAVGSQDLFTASLPPEKNCQDGLCEGDLGVTLSFGSLQTLTVG 793
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPLDLSYRKVSTLQNRQSRQWRLACESASSTEV 840
QY 794 SSLELVIVITVMNAGEDSYGTWSLYYPAGLSHRVSGAQKPHQSALRLACETV-PTED 852
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTFF 900
QY 853 EG-LRSSRCNVNHPHFEGSNGTIVTFDYSKATLDGRMLMRASSENKASSSKATF 911
QY 901 QLELPVKYAVVMVYTSHGVSSTKYNF-TASENTSRVMQHOYQVSNLQORSUPISLVFLVP 959
QY 912 QLELPVKYAVVTMISROEESTKYENFATSDCKMKEAHRVYRVNLSQORDLAISINFVWP 971
QY 960 VRLNQTVIWDPPQVTFSENLSSTCHTERLPSSHDSDFLAELRKAPVNVCSIAVCQICDI 1019
QY 972 VLLNGVAVMDVVMAPSQSL--PCVSEKPKPHSDFLTQISRSPLMDCSIADCLQFRCDV 1029
QY 1020 PFFGIQBEFNATLKGNLISFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTE 1079
QY 1030 PSFSVQBELDFTLKGNLISFGWVRETLOKVLVVSVAEITFTDTSVYSQLPGQEAFFRAQME 1089
QY 1080 TKVBPPEVPNPPLPIVIGSSVGGLLLLALITAAALYKLGFFKROYKDMSE 1128
QY 1090 MVLEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRYKEMLEED 1138

```

Search completed: November 25, 2003, 14:16:34
Job time : 40.2978 secs


```
QY 61 RLOQVPEAVNNMGLSLAATSPPOLLAGCPTVHTQCSNTYVVKGLCFPGSNLRQOPQK 120
DB 77 RLOQVPEAVNNMGLSLAATSPPOLLAGCPTVHTQCSNTYVVKGLCFPGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMQYSSEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMQYSSEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELNTVASKPRDRHVQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELNTIASKPRDRHVQIN 316
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMQSGEFSAAITSGNPLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMQSGEFSAAITSGNPLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKGTOIGAYFCASICSVDVDSNGSTDVLIGAPHYYEQTGGOVSVCLP 480
DB 437 QNTGWESNANVKGTOIGAYFCASICSVDVDSNGSTDVLIGAPHYYEQTGGOVSVCLP 496
QY 481 PRGQARWQCDVAVLGEQOPNGRFGAALTVLGDVNGDKLTDAVATGAPGEEDNRGAVLYF 540
DB 497 PRGQARWQCDVAVLGEQOPNGRFGAALTVLGDVNGDKLTDAVATGAPGEEDNRGAVLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIINEFNPVARNVFCNDQVVKGEAVRVCVHVOKSTRDLREGQIQSVVT 660
DB 617 PVLVRKAIINEFNPVARNVFCNDQVVKGEAVRVCVHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDLSGRPHSRVAFNETKSTRTOVLGLTQTCETLKLQPCIEDPVPSPVLRNLF 720
DB 677 YDLALDLSGRPHSRVAFNETKSTRTOVLGLTQTCETLKLQPCIEDPVPSPVLRNLF 736
QY 721 SLVGTPLSAFNLRLPVLAEDAQLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFNLRLPVLAEDAQLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPFDLDSYRKVSTLONORSORSHWLACESASSTEV 840
DB 797 GPREFNVTVTVRNDGEDSYRTQVTFPFDLDSYRKVSTLONORSORSHWLACESASSTEV 856
QY 841 SGALKSTCSINHPIPPENSEVFTNITFDVDSKASLGNKLLKANVTSENMPRNTKTEF 900
DB 857 SGALKSTCSINHPIPPENSEVFTNITFDVDSKASLGNKLLKANVTSENMPRNTKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNCSTAVCQRTQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNCSTAVCQRTQCDIP 1036
QY 1021 FFGIOEFNATLKGNLSDFWYIKTSHNHLIIYSTAILFNDVSTFLLPQCGAFVRSQTEF 1080
DB 1037 FFGIOEFNATLKGNLSDFWYIKTSHNHLIIYSTAILFNDVSTFLLPQCGAFVRSQTEF 1096
QY 1081 KVEPFEPVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1137
DB 1097 KVEPFEPVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1153
```

```
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286.889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3
Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLOTENAMTFOENARGFQSVVOLQGSRVVVGAPQEIIVAAANORGSLYQCDYSTGSCCEPI 60
DB 17 FNLOTENAMTFOENARGFQSVVOLQGSRVVVGAPQEIIVAAANORGSLYQCDYSTGSCCEPI 76
QY 61 RLOQVPEAVNNMGLSLAATSPPOLLAGCPTVHTQCSNTYVVKGLCFPGSNLRQOPQK 120
DB 77 RLOQVPEAVNNMGLSLAATSPPOLLAGCPTVHTQCSNTYVVKGLCFPGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMQYSSEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMQYSSEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELNTVASKPRDRHVQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELNTIASKPRDRHVQIN 316
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMQSGEFSAAITSGNPLSTVGSYDWAG 360
DB 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMQSGEFSAAITSGNPLSTVGSYDWAG 376
```


Qy	361	GVPFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAAILLRNRVQSILVLGAPRYQHIGLVAMFR	420
Dd	377	GVPFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAAILLRNRVQSILVLGAPRYQHIGLVAMFR	436
Qy	421	QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSDDLVLIGAPHYYEQTRGGQSVCP	480
Dd	437	QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSDDLVLIGAPHYYEQTRGGQSVCP	496
Qy	481	PRGORARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAI GAPGEEDNRGAVLYLF	540
Dd	497	PRGORARWQCDAVLYGEQGQPWGRFGAALTVLGDVNGDKLTDVAI GAPGEEDNRGAVLYLF	556
Qy	541	HGTSGSGISPSHSQRAGSKSLPRLQYFGOSLSGGQDLTMDGLVDLTVGAGQHVVLLRSQ	600
Dd	557	HGTSGSGISPSHSQRAGSKSLPRLQYFGOSLSGGQDLTMDGLVDLTVGAGQHVVLLRSQ	616
Qy	601	PVLVRVKAIMEFNPREVARNVPECNQOVVKGKEAGEVRVCLHVOKSTRDRLREQIOSVVT	660
Dd	617	PVLVRVKAIMEFNPREVARNVPECNQOVVKGKEAGEVRVCLHVOKSTRDRLREQIOSVVT	676
Qy	661	YDIALDSGRPHSRVAFNETKNSTRTOVLGLTOTCETLKLQIPNCIEDDPVSPIVLRLNF	720
Dd	677	YDIALDSGRPHSRVAFNETKNSTRTOVLGLTOTCETLKLQIPNCIEDDPVSPIVLRLNF	736
Qy	721	SLVGTPLSAFCNLRPVLAEDAORLFTALPFPEKXCGNDNICQDDLSITTSFMSLDCLVWG	780
Dd	737	SLVGTPLSAFCNLRPVLAEDAORLFTALPFPEKXCGNDNICQDDLSITTSFMSLDCLVWG	796
Qy	781	GPREFNVTVTRNDGEDSYRTQTFFFPDLDSYRKYSTLQNORSQRSWRLACESASTEV	840
Dd	797	GPREFNVTVTRNDGEDSYRTQTFFFPDLDSYRKYSTLQNORSQRSWRLACESASTEV	856
Qy	841	SGALKSTSCSINHPIPPENSESEVTNITFDVDSKASLGKULLKANVTSENMPRTNKTEF	900
Dd	857	SGALKSTSCSINHPIPPENSESEVTNITFDVDSKASLGKULLKANVTSENMPRTNKTEF	916
Qy	901	OLELPVKYAYVMVMTSHGVSTKYLNFTASENTRVMQHOYOVNGLGORSUPI SILVFLVPV	960
Dd	917	OLELPVKYAYVMVMTSHGVSTKYLNFTASENTRVMQHOYOVNGLGORSUPI SILVFLVPV	976
Qy	961	RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORI QCDIP	1020
Dd	977	RLNOTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNCSIAVCORI QCDIP	1036
Qy	1021	FFGIQEEFNATLKGNLSFDWIYIKTSHNHLLIVSTAEIFLNDVSFTLLPGOGAFVRSQTET	1080
Dd	1037	FFGIQEEFNATLKGNLSFDWIYIKTSHNHLLIVSTAEIFLNDVSFTLLPGOGAFVRSQTET	1096
Qy	1081	KVBPFEPVNPULPIVGSVUGLLILLALIITAALYKLGFFKFOYKDMMEGGPPGAEPQ	1137
Dd	1097	KVBPFEPVNPULPIVGSVUGLLILLALIITAALYKLGFFKFOYKDMMEGGPPGAEPQ	1153

Db 557 HGTSGGSPSHSQRASGLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 616
Qy 601 PVLRYKAIEMFNPREVARNVFNCDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRYKAIEMFNPREVARNVFNCDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVLRNLF 720
Db 677 YDLALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVLRNLF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPEKKNCGNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDCGDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDCGDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 856
Qy 841 SCALKSTSCSINHPIPPENSEVTENITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 857 SCALKSTSCSINHPIPPENSEVTENITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSPLSLVFLVPV 976
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRTQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRTQCDIP 1036
Qy 1021 PFGIOBEFNATLKGNSLDWYIKTSHNHLIIIVSTAEILLFNDVSFTLLPGQAFVRSQTEF 1080
Db 1037 PFGIOBEFNATLKGNSLDWYIKTSHNHLIIIVSTAEILLFNDVSFTLLPGQAFVRSQTEF 1096
Qy 1081 KYPEVEPNPLIIVGSSVGGLLALLALITAAALYKLGFFKRYQKDWMSGGPPGABPQ 1137
Db 1097 KYPEVEPNPLIIVGSSVGGLLALLALITAAALYKLGFFKRYQKDWMSGGPPGABPQ 1153

RESULT 4.

US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/362,652
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
FILING DATE: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
FILING DATE: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 278666/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGOSVVOLOQSRVVGGAPOEIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGOSVVOLOQSRVVGGAPOEIVAAORGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMVSLGLSLAATTPPQLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMVSLGLSLAATTPPQLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQPOK 136
Qy 121 PPEALRCQPEDSDIAFLVDGSGSIIIPDFRRAKEPISTVMEQLKSKTFLSLMOYSEEF 180
Db 137 PPEALRCQPEDSDIAFLVDGSGSIIIPDFRRAKEPISTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHTTFEFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFINITGARKNAFKILILI 240
Db 197 RIHTTFEFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFINITGARKNAFKILIVI 256
Qy 241 TDGKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSKQELNLTVASKPPRDHVFQIN 300
Db 257 TDGKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSKQELNLTIASKPPRDHVFQIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITNSGPELLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITNSGPELLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDMDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLLVLIGAPHYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLLVLIGAPHYEQTRGGQVSVCP 496
Qy 481 PRQQRARWQCDAVLYGEOGQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQQRARWQCDAVLYGEOGQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGGSPSHSQRASGLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 600
Db 557 HGTSGGSPSHSQRASGLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 616
Qy 601 PVLRYKAIEMFNPREVARNVFNCDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRYKAIEMFNPREVARNVFNCDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVLRNLF 720
Db 677 YDLALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVLRNLF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPEKKNCGNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDCGDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLVSRVKVSTLQNRQSRWRLACSSASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 976
Qy 961 RLNQTVIWDPRPOVTFSENLSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP 1020
Db 977 RLNQTVIWDPRPOVTFSENLSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP 1036
Qy 1021 FFGIOEFNATLKNLSFDMYIKTSHNHLIVSTAEILFNDVFTLLPGOGAFVRSOTET 1080
Db 1037 FFGIOEFNATLKNLSFDMYIKTSHNHLIVSTAEILFNDVFTLLPGOGAFVRSOTET 1096
Qy 1081 KVBPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKXDMMSGEGPPGAEPQ 1137
Db 1097 KVBPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKXDMMSGEGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,869
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCPEI 76
Qy 61 RLQOPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGICFLFGSLNRQOPQ 120
Db 77 RLQOPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTYVKGICFLFGSLNRQOPQ 136
Qy 121 FPEALRGCCQEDSDIAFLVDGSGSIIPHDFRAKEPISTVMEOLKSKTLLFSLMQSSEF 180
Db 137 FPEALRGCCQEDSDIAFLVDGSGSIIPHDFRAKEPISTVMEOLKSKTLLFSLMQSSEF 196
Qy 181 RIHFTPEKQNNPNRSLIKPIITQLLGRTHATGIRKVVRELFININGARKNAFKILILI 240
Db 197 RIHFTPEKQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFININGARKNAFKILVVI 256
Qy 241 TDGEKEGDPGLGYEDVPEADREGVIRYVIGVDAPFSEKSRQELNTVASKPRDHVQIN 300
Db 257 TDGEKEGDPGLGYEDVPEADREGVIRYVIGVDAPFSEKSRQELNTVASKPRDHVQIN 316
Qy 301 NFEALXTIQNOLREKIFAIEGTQSGSSSEFHEMSQEGFSAITNSGPLLLSTVGSYDWAG 360
Db 317 NFEALXTIQNOLREKIFAIEGTQSGSSSEFHEMSQEGFSAITNSGPLLLSTVGSYDWAG 376
Qy 361 GVPLYTSKEKSTFINMTRVDSMDNAYLGAAAAIILNRNVQSLVLGAPRYQHIGLVAMPR 420
Db 377 GVPLYTSKEKSTFINMTRVDSMDNAYLGAAAAIILNRNVQSLVLGAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDILVIGAPHYIEQTRGGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDILVIGAPHYIEQTRGGQSVCP 496
Qy 481 PRGORARWOCDAVLGYEQGQPGWRFGAALTVDLGVNGDKLTDVAICAPGEEDNRGAYVLF 540
Db 497 PRGORARWOCDAVLGYEQGQPGWRFGAALTVDLGVNGDKLTDVAICAPGEEDNRGAYVLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLRVKAIWEPNPREVARNVFECNDQVVKGEAGEVVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 617 PVLRVKAIWEPNPREVARNVFECNDQVVKGEAGEVVRVCLHVQKSTRDRLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVFNENKSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPILVRLNF 720
Db 677 YDLALDSGRPHSRVFNENKSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPILVRLNF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEDAORLFTALPPFEKNKCGNDNICDDLSITTFMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAORLFTALPPFEKNKCGNDNICDDLSITTFMSLDCLVWG 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLVSRVKVSTLQNRQSRWRLACSSASSTEV 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLVSRVKVSTLQNRQSRWRLACSSASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 976
Qy 961 RLNQTVIWDPRPOVTFSENLSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP 1020
Db 977 RLNQTVIWDPRPOVTFSENLSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP 1036

Qy	1021	FFGIQEBFNATLKNLSFDMYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET	1080
Db	1037	FFGIQEBFNATLKNLSFDMYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET	1096
Qy	1081	KVEPFEVNPPLPLIVGSSVGLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ	1137
Db	1097	KVEPFEVNPPLPLIVGSSVGLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ	1153
RESULT 6			
US-08-482-293A-3			
; Sequence 3, Application US/08482293A			
; Patent No. 5831029			
; GENERAL INFORMATION:			
; APPLICANT: Gallatin, W. Michael			
; APPLICANT: Van der Vieren, Monica			
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit			
; NUMBER OF SEQUENCES: 103			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun			
; STREET: 233 South Wacker Drive, 6300 Sear Tower			
; CITY: Chicago			
; STATE: Illinois			
; COUNTRY: United States			
; ZIP: 60606-6402			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/482,293A			
; FILING DATE:			
; CLASSIFICATION: 530			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/173,497			
; FILING DATE: 23-DEC-1993			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/286,889			
; FILING DATE: 5-AUG-1994			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/362,652			
; FILING DATE: 21-DEC-1994			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Williams Jr., Joseph A.			
; REGISTRATION NUMBER: 38,659			
; REFERENCE/DOCKET NUMBER: 27866/32684			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 312-474-6300			
; TELEFAX: 312-474-0448			
; TELEX: 25-3856			
; INFORMATION FOR SEQ ID NO: 3:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1153 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-482-293A-3			
Query Match 99.8%; Score 5862; DB 2; Length 1153;			
Best Local Similarity 99.3%; Pred. No. 0;			
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ENLDTENAMTFQENARGFGQSVVQLOGSRVVVQCAPQEIIVAAVORGLSYOCDSYTGSCPEI	60
Db	17	ENLDTENAMTFQENARGFGQSVVQLOGSRVVVQCAPQEIIVAAVORGLSYOCDSYTGSCPEI	76
Qy	61	RLQVPVAVNMVSLGLSLAATTPSPQLIACGPTVHQTCSENTYVKGICFLFGSNLRQOPQK	120
Db	77	RLQVPVAVNMVSLGLSLAATTPSPQLIACGPTVHQTCSENTYVKGICFLFGSNLRQOPQK	136
Qy	121	FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKXSKTLFSLMQYSEEP	180

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

```

; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817478e1 Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 127866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-943-363-3

Query Match          99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY      1  FNLDTENAMTFQBNARGFGQSVVQLGSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB      17  FNLDTENAMTFQBNARGFGQSVVQLGSRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY      61  RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPK 120
DB      77  RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPK 136

QY      121  FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLKSKTLFSLMQYSSEF 180
DB      137  FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLKSKTLFSLMQYSSEF 196

QY      181  RIHFTPFQNNPNPSLKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKLLI 240
DB      197  RIHFTPFQNNPNPSLKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKLLI 256

QY      241  TDGEKFGDPLGVEDVPIPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPRDHVFQIN 300
DB      257  TDGEKFGDPLGVEDVPIPEADREGVIRYVIGVDGDAFSEKSRQELNTVASKPRDHVFQIN 316

QY      301  NFEALKTIONQREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG 360
DB      317  NFEALKTIONQREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG 376

```

```

QY      361  GVFLVTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVOSLVLGAPRYQHIGLVAMFR 420
DB      377  GVFLVTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVOSLVLGAPRYQHIGLVAMFR 436

QY      421  QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDLVLIGAPHYYEQTGRGOVSVCPL 480
DB      437  QNTGMWESNANVKGTQIGAYFGASICSVDVDSNGSTDLVLIGAPHYYEQTGRGOVSVCPL 496

QY      481  PRGQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDAVTCAPCEEDNRGAVYLF 540
DB      497  PRGQARWQCDVAVLYGEOQPMGRFGAALTVLGDVNGDKLTDAVTCAPCEEDNRGAVYLF 556

QY      541  HGTSGSISPSHSORIASKJSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLLRSQ 600
DB      557  HGTSGSISPSHSORIASKJSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

QY      601  PVLRRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
DB      617  PVLRRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

QY      661  YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPPVSPVLRNLF 720
DB      677  YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPPVSPVLRNLF 736

QY      721  SLVGTPLSAFGNLRPVLAEQAQLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780
DB      737  SLVGTPLSAFGNLRPVLAEQAQLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796

QY      781  GPREFNVTVVRNDEGDSYRQVTFPFLDLISYRKVSTLQNRORSORSWRLACESASTEV 840
DB      797  GPREFNVTVVRNDEGDSYRQVTFPFLDLISYRKVSTLQNRORSORSWRLACESASTEV 856

QY      841  SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
DB      857  SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 916

QY      901  QLEPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHQYQVSNLQGRSLPISLVFLVPV 960
DB      917  QLEPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHQYQVSNLQGRSLPISLVFLVPV 976

QY      961  RLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQIQCDIP 1020
DB      977  RLNQTVIWDROPVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQIQCDIP 1036

QY      1021  PFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQCGAFVRSQDET 1080
DB      1037  PFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQCGAFVRSQDET 1096

QY      1081  KVEPPEVPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMMSGGPPGABPQ 1137
DB      1097  KVEPPEVPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMMSGGPPGABPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER FILING DATE: 1993-12-23
; EARLIER FILING DATE: 1994-08-05
; EARLIER FILING DATE: 1994-08-05
; EARLIER FILING DATE: 1994-12-21
; EARLIER FILING DATE: 1994-12-21
; EARLIER FILING DATE: 1997-10-03

```

; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 99.8%; Score 5862; DB 3; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGOSVVOLOGSRVVGAPQEIIVAAVNOGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGOSVVOLOGSRVVGAPQEIIVAAVNOGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCTSENTYVKGCLFGLGSLNRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCTSENTYVKGCLFGLGSLNRQOPQK 136

Qy 121 FPEALRCPCQEDSDIAFLIDGSGSIIPHDFFRAKEFISTVMEQLKKSXTLPSLMQYSEEF 180
Db 137 FPEALRCPCQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKKSXTLPSLMQYSEEF 196

Qy 181 RIHFTTFKEFQNNPNRSLIKPITQLGRTHATGIRKVVRELFNITGARKNAFKILILI 240
Db 197 RIHFTTFKEFQNNPNRSLIKPITQLGRTHATGIRKVVRELFNITGARKNAFKILIVI 256

Qy 241 TDGKFKGDPGLGYEDVPEADREGVIRVVGVDGAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGKFKGDPGLGYEDVPEADREGVIRVVGVDGAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFNMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOCHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFNMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOCHIGLVAMFR 436

Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYTEQTRGGQVSCPL 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYTEQTRGGQVSCPL 496

Qy 481 PRGQARWQCDVLYGEGQPMGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQPMGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDLTMDGLVDLTGAGHVLILLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLILLRSQ 616

Qy 601 PVLRVKATMEFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQTSVVT 660
Db 617 PVLRVKATMEFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQTSVVT 676

Qy 661 YDLALDSGRPHSRAVNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVPSPVLRNLF 720
Db 677 YDLALDSGRPHSRAVNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVPSPVLRNLF 736

Qy 721 SLVGTPLSAFGNLRPLVLAEDAQRLLTALPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRPLVLAEDAQRLLTALPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796

Qy 781 GREFNVTVVRNDCGDSVRTQVTFPFLDLRYKVTSLQNSORSWRLACESASTEV 840
Db 797 GREFNVTVVRNDCGDSVRTQVTFPFLDLRYKVTSLQNSORSWRLACESASTEV 856

Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916

Qy 991 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVWQHOYQVSNLQORSLPISLFLVPV 960

RESULT 9

US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 99.8%; Score 5862; DB 4; Length 1153;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGOSVVOLOGSRVVGAPQEIIVAAVNOGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGOSVVOLOGSRVVGAPQEIIVAAVNOGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCTSENTYVKGCLFGLGSLNRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCTSENTYVKGCLFGLGSLNRQOPQK 136

Qy 121 FPEALRCPCQEDSDIAFLIDGSGSIIPHDFFRAKEFISTVMEQLKKSXTLPSLMQYSEEF 180
Db 137 FPEALRCPCQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKKSXTLPSLMQYSEEF 196

Qy 181 RIHFTTFKEFQNNPNRSLIKPITQLGRTHATGIRKVVRELFNITGARKNAFKILILI 240
Db 197 RIHFTTFKEFQNNPNRSLIKPITQLGRTHATGIRKVVRELFNITGARKNAFKILIVI 256

Qy 241 TDGKFKGDPGLGYEDVPEADREGVIRVVGVDGAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGKFKGDPGLGYEDVPEADREGVIRVVGVDGAFRSEKSRQELNTIASKPPRDHVFQVN 316

Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360

Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGGOVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGGOVSCPL 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAAGHVLRLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAAGHVLRLRSQ 616
Qy 601 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPSRAVENETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 720
Db 677 YDLALDSGRPSRAVENETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPEKKNCGNDNICQDDLSITFFSMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPEKKNCGNDNICQDDLSITFFSMSLDCLVWG 796
Qy 781 GPREFNVTVRNDEGSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVRNDEGSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVFLVPV 976
Qy 961 RLNOTVIWDRPOVTFSENLSCTHTKERLPKSHDFLAELRKAPVNVCSIAVCRIQCDIP 1020
Db 977 RLNOTVIWDRPOVTFSENLSCTHTKERLPKSHDFLAELRKAPVNVCSIAVCRIQCDIP 1036
Qy 1021 FFGQEEFNATLKGNSLSDWIKTSHNHLIIVSTAELFNDVSFTLLPGQAFVRSOTET 1080
Db 1037 FFGQEEFNATLKGNSLSDWIKTSHNHLIIVSTAELFNDVSFTLLPGQAFVRSOTET 1096
Qy 1081 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMMSGGPPGABPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMMSGGPPGABPQ 1153

RESULT 10

US-08-476-062A-43

; Sequence 43, Application US/08476062A

; Patent No. 5877275

; GENERAL INFORMATION:

; APPLICANT: Arnaout, M. Amin

; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY

; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476, 062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216, 081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637, 830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539, 842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212, 573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29, 066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-476-062A-43

Query Match 99.2%; Score 5831.5; DB 2; Length 1152;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGTPVHOTCSENTYVKGLCFLFGSNLURQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGTPVHOTCSENTYVKGLCFLFGSNLURQPOK 136
Qy 121 FPEALRCPCQEDSDIAFLIDSGSGIIIPHDFFRAKEFISTVMEQLKSKTLFSLMQYSEEF 180
Db 137 FPEALRCPCQEDSDIAFLIDSGSGIIIPHDFFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLRPIITQLLORTHTATGIRKVVRELFNITNGARKNAPKILVI 240
Db 197 RIHFTFKFQNNPNRSLRPIITQLLORTHTATGIRKVVRELFNITNGARKNAPKILVI 256
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQIN 316
Qy 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGGOVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGGOVSCPL 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAAGHVLRLRSQ 600

556 HGTSGGSPSHSQRISAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRSQ 615
601 PVLRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
616 PVLRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
661 YDLALDSGRPHSRVAFNETKSTRTOVLGLTQTCETLKLQPNCEIDPVPVLRNLF 720
676 YDLALDSGRPHSRVAFNETKSTRTOVLGLTQTCETLKLQPNCEIDPVPVLRNLF 735
721 SLVGTPLSAFAGNLRPLVLAEDAQRFTALFPFKEKNCNDNICODDLSITFSFMSLCLVVG 780
736 SLVGTPLSAFAGNLRPLVLAEDAQRFTALFPFKEKNCNDNICODDLSITFSFMSLCLVVG 795
781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
796 GPRESNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
901 QLELPVKYAVYVWVTSHGVTYKLNFTASNTSRVWQHQQVNSLQGRSLPISLVFLVPV 960
916 QLELPVKYAVYVWVTSHGVTYKLNFTASNTSRVWQHQQVNSLQGRSLPISLVFLVPV 975
961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVWNCSTAVCQRIQCDIP 1020
976 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVWNCSTAVCQRIQCDIP 1035
1021 FFGIOBEFNATLKNLSFDWYIKTSINHLLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
1036 FFGIOBEFNATLKNLSFDWYIKTSINHLLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1095
1081 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKDMSEGGPPGAEPO 1137
1096 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKDMSEGGPPGAEPO 1152

RESULT 11

PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29, 066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906

TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43

Query Match 99.2%; Score 5831.5; DB 5; Length 1152;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOBIVAAANQSGSLYQCDYSTGSCSEPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOBIVAAANQSGSLYQCDYSTGSCSEPI 76
QY 61 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYKGLCFGLGSLNRLOQPOK 120
DB 77 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYKGLCFGLGSLNRLOQPOK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPRRAKEFIPTVMEQLKKSKTLFSLMOYSBEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPRRAKEFIPTVMEQLKKSKTLFSLMOYSBEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLGRTHRTATGIRKVVRELFINITNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLGRTHRTATGIRKVVRELFINITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPRDRHVQIN 316
QY 301 NFEALKTITONOLREKIPIAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDNAG 360
DB 317 NFEALKTITONOLREKIPIAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDNAG 376
QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQGVSVCLP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQGVSVCLP 496
QY 481 PRGQARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDAIAGAPEEDNRGAVYLF 540
DB 497 PRG-RARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDAIAGAPEEDNRGAVYLF 555
QY 541 HGTSGGSPSHSQRISAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRSQ 600
DB 556 HGTSGGSPSHSQRISAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRSQ 615
QY 601 PVLRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 616 PVLRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKSTRTOVLGLTQTCETLKLQPNCEIDPVPVLRNLF 720
DB 676 YDLALDSGRPHSRVAFNETKSTRTOVLGLTQTCETLKLQPNCEIDPVPVLRNLF 735
QY 721 SLVGTPLSAFAGNLRPLVLAEDAQRFTALFPFKEKNCNDNICODDLSITFSFMSLCLVVG 780
DB 736 SLVGTPLSAFAGNLRPLVLAEDAQRFTALFPFKEKNCNDNICODDLSITFSFMSLCLVVG 795
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 796 GPRESNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915


```
Qy 901 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
Db 916 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 975
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDIP 1020
Db 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDIP 1035
Qy 1021 FFGIOBEFNATLKNLSFDWIKYKTSNHNLLIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
Db 1036 FFGIOBEFNATLKNLSFDWIKYKTSNHNLLIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGPPGAEPQ 1152

RESULT 12
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO:2
; LENGTH: 1152
5424399-2

Query Match 99.28; Score 5831.5; DB 6; Length 1152;
Best Local Similarity 99.08; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQIIVAAVQORGSILYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQIIVAAVQORGSILYQCDYSTGSCPEI 76

Qy 61 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYKVGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYKVGCLFLFGSNLRQOPQK 136

Qy 121 FPEALRGCPQEDSDIAPFDVSGSIIIPDFRRAKEFISTVMEQLKSKTILFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAPFDVSGSIIIPDFRRAKEFISTVMEQLKSKTILFSLMQYSEEF 196

Qy 181 RIHFTFKFQNNPRSLIPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPRSLIPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

Qy 241 TDGEKFGDPLGYEDVIEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHFQIN 300
Db 257 TDGEKFGDPLGYEDVIEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHFQVN 316

Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHENSQEGFAAITSNGPLLSVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHENSQEGFAAITSNGPLLSVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINMTFRVSDMDNDAYLGVAIAIILNRVOSULVGLAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTFRVSDMDNDAYLGVAIAIILNRVOSULVGLAPRYOHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTQIGAVFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAVFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQVSCPL 496

Qy 481 PRGQARWQCDAYLYGQGOQPWGRFGAALTVLGQVNGDKLTDVAIGAPGEDNRRGAYVLF 540
|||
```

```
Db 497 PRG-RARWQCDAYLYGQGOQPWGRFGAALTVLGQVNGDKLTDVAIGAPGEDNRRGAYVLF 555
Qy 541 HGTSGGIGSPSHSQRISAGSKLSPLQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ 600
Db 556 HGTSGGIGSPSHSQRISAGSKLSPLQYFGQSLSGGQDLTMDGLVDLITVGAQGHVLLRSQ 615
Qy 601 PVLVRKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
Db 616 PVLVRKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 675
Qy 661 YDLALDSGPHSRVAVNETKNSRROTQVLTOTCETLKLQLPNCIEDPVSIVLRLNF 720
Db 676 YDLALDSGPHSRVAVNETKNSRROTQVLTOTCETLKLQLPNCIEDPVSIVLRLNF 735
Qy 721 SLVGTPLSAFNGNLRPVLAEADAQRLFTALPFPFENKNCNDNICODDLSITFSFMSLDCLVG 780
Db 736 SLVGTPLSAFNGNLRPVLAEADAQRLFTALPFPFENKNCNDNICODDLSITFSFMSLDCLVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
Db 796 GPRESNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPRTNKTEF 915
Qy 901 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
Db 916 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 975
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDIP 1020
Db 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCORIQCDIP 1035
Qy 1021 FFGIOBEFNATLKNLSFDWIKYKTSNHNLLIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
Db 1036 FFGIOBEFNATLKNLSFDWIKYKTSNHNLLIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGPPGAEPQ 1152

RESULT 13
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arraout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
```

FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-062A-44

Query Match 59.0%; Score 3469; DB 2; Length 1163;
Best Local Similarity 61.0%; Pred. No. 7,1e-284;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFQSVQLQGSRRVVVGAPOEIVAAVNRQSLQCYDSTGSCBPI 60
DB 20 FNLDTTEELTAFRVDSAGFSDSVVQVANSWVVGAPQKITAANQOTGGLYQCYDSTGACBPI 79

QY 61 RLQVPEAVNMLGLSLAATTPPOLLACGPTVHTCSENTYVKGICLFLGSLNLRQOQOK 120
DB 80 GLQVPEAVNMLGLSLASTTSPQLLACGPTVHHECGRMVYLTGLCFLLGPT--OLTOR 137

QY 121 FPEALRGCPQEDSLAFLVDGSGSIIPHDFRAKEFIPTVMEQLKSKTLFSLMOYSBEF 180
DB 138 LPVDSQECRQODIVFLIDGSGSISSRNFAFMFVRAVISQFORPSTQFSLMQFSNK 197

QY 181 RIHFTKPEQNNPNSRIKPTQLGRTHATGINKVRELFNITNGARKNAFKILILI 240
DB 198 QTHFTFEPRRTSNPLSLASVHQLQGYTTATQNVVHRLPHFASYGARRDATKILIVI 257

QY 241 TDEKFGDPLGVEDVPEADRGVIRYVIGVDAPRSEKSRQELNNTVASKPPRDHVFQIN 300
DB 258 TDGKEGSLDYKDVIPNADAAGIIRYAIGVGLAFQNRNSKELNDIASKPSQEHIFKVE 317

QY 301 NFEALKTIONQIREKIFATEGTQTSSEFHEMSQEGFSAAITNSGPLLSTVGSVDWAG 360
DB 318 DFDAUKDIONQIREKIFATEGTTSSEFHEMSQEGFSAVFTPDGVLGAVGFTWSG 377

QY 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAILRNVRQSLVLGAPRYQHIGLVAMFR 420
DB 378 GAFLYPPNNSPTFINMSQENVMDRDSYLGSTELALWKGVSILVGLGAPRYQHTKAVIFT 437

QY 421 QNTGHWESNANKVGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSVCLP 480
DB 438 QVSRQWRMKAENVGTQIGSYFGASLCSVDVDTGSTDVLIGAPHYYEOTRGQGVSVCLP 497

QY 481 PRGORARWOCDAVLVGEQCPWGRFCAALTVLGDVNGDKLTDAVIGAPCEENRGAVALYL 540
DB 498 PRGWR-RWNCDAVLVGEQCPWGRFCAALTVLGDVNGDKLTDAVIGAPCEENRGAVALYL 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSLSGGQDLTMDGLVDLTVAQOQHVLRLRQ 600
DB 557 HGVLGFSISPSHSQRIAGSKLSPRLQYFQSLSGGQDLTMDGLVDLAVGARGQVLLRLR 616

QY 601 PVLRVKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIOSVT 660
DB 617 PVLWVGVSQFIPAEIPRSAFECEQVSEQLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

QY 661 YDLALDPSRPHSRVAVNETKNSRRROTQVLGTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
DB 677 LDIALDPSRPHSRVAVNETKNSRRROTQVLGTQTCETLKLQLPNCIEDPVPSPVILRLNF 736

QY 721 SLVGTPLSAGNLRPVLAEDAQRFLTALFPFEKNCNDNICODDLSTIFSPMSLDCLVVG 780

DB 737 TLVGKPLLAFLRNLRPMJLAALAQRYFTASLPFKCNKGADHICQDNLGIFSFPFGKSLLVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSTYRKVSTLQNRQSRQWRMLACESASSTEV 840
DB 797 SNLENAEVMVNDGEDSYGTTITFSPAGLSYRYVAGQXQGLRSLHLTCDAPVG-- 854
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
DB 855 SQGTWSTSCRINHLLIFRGAQITELATFDVSPKAVLGDRLILLTANVSSENNTPTSTKTF 914
QY 901 QLELPVKYAVYVTVSHGVSTKYLNFAS-ENTSRVMQHOYOVSNLGRSIPISLVFLVP 959
DB 915 QLELPVKYAVYVTVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
QY 960 VRLNQTVIWDROPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORIQCDI 1019
DB 975 VELNQEAVMVDVEVSHQNPNSLRCSSEKIAAPPADFIAHQKNPVLDCSIAGCLRCFCDV 1034
QY 1020 PFFGQIEFNATLKNLSFDWYIKTSHNHLIVTAEILFNDSVFTLLPGGQAFVRSOTE 1079
DB 1035 PSFSVQBELDFTLKNLSFGWVRQILQKQVSVSWVAEITFTSVYSQLPGQAFPMRAQTT 1094
QY 1080 TKVEPFEVNPPLPLIVGSSVCGLLALLITALYKLGFFKROKYMHEE 1128
DB 1095 TVLEKYKVHNPPLIVGSSIGGLLLALLITAVLYKVGFFKROKYMHEE 1143

RESULT 14
PCT-US96-01314-44
Sequence 44, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
TITLE OF INVENTION: ANTAGONISTS
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 1163
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-44

Query Match 59.0%; Score 3469; DB 5; Length 1163;
Best Local Similarity 61.0%; Pred. No. 7,1e-284;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

[illegible]

Db 318 DFDALKIQNOLKEKIFAIBGTETETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377
QY 361 GVFLYTSKEKSTFINMTFRVDSMDMDAYLGVAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 378 CAFLYPPNMSFTFINMSQENVDMRDSYLGSTELALWKGVQSLVGLGAPRYQHIGKAVIFI 437
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYEQTRGGQVSVCP 497
QY 481 PRGQARWQCDAYLGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 498 PRGMR-RWMCDAVLGEGHPWGRFGAALTVLGDVNGDKLTDVIGAPGEENRGAVYLF 556
QY 541 HGTSGSISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLGPSISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRYKAIMEFNPREVANRPECNDQVKGKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 617 PVLWGVSMQFIPASIEPRSAFECRQVYSEQTLVQSNICLYIDKSKNLLGSRDLQSSVT 676
QY 661 YDLALDSCRPHSRVAFNETKSTRTQTVLGLTQTCETLKLQPCIEDPVSPVILRLNF 720
Db 677 LDALAPGRLSPRAFQETKNRSLSRVRVLGKAHCENFNLLPSCVEDSVIPIILRLNF 736
QY 721 SLVGTPLGAFGLRPLVAEDAQRLETFALPFEKNCNDNICODDLSITFSFMSLCLVVG 780
Db 737 TLVGPLLAFLNRLPMLAALQRYFTASLPFEKNCADHICODNLGISFSFPLGLKSLVG 796
QY 781 GPREFNVTVVRDGEDSVRTQVTFPFLDLSYRKVSTLQNRORSORSWRLACESASSTEV 840
Db 797 SNLELNAEVMWMDGEDSVYGTITFESHAPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENMMPTNKTEF 900
Db 855 SQGTWSTSCRINHLPFRGAQITFLATFDVSPRAVGLDRLLIANVSSENNIPRTSKTIF 914
QY 901 QLELPVKYAVYVVTSHGVSTKYLNFTAS-ENTSRVMOHQYQVSNLQORSPLISLVFLVP 959
Db 915 QLELPVKYAVYVVSHEQFTKYNFSESEKESHVAMHRYQVNNLQORDLPVSINFVWP 974
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKELPSHSDFLAELRKAPVNVNCSIAVCQRIQCDI 1019
Db 975 VELNQAAMVMDVEVSHQNPRLSCSEKIAPPASDFLAHQKNPVLDCSIAGCLFRCDV 1034
QY 1020 PFTGQEEFNATLKGNLSPDWYIKTSHNHLIIYSTAEILLFNDVSFTLLPCQGAFFVSQTE 1079
Db 1035 PSFSVQBEELFTLKGNLSPGWVRQIILQKKVSVVSVAEIIPDTSVVSQLPQGEAFMRAQTI 1094
QY 1080 TKVEPPEVPNPLPIVSSVGGLLLALITALYKLGFFKRYQKDMWSE 1128
Db 1095 TVLEKYKVNHPILIVGSSIGGLLLALITALYKVGFFKRYQKEMWEE 1143

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:12:29 ; Search time 14.9579 Seconds
(without alignments)
7310.102 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMVSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5868	99.9	1153	1	RWHU1B
2	4476	76.2	1153	2	S00551
3	3483	59.3	1163	1	RWHU1C
4	1548.5	26.4	1170	2	S03308
5	1533.5	26.1	1163	2	I56126
6	1149	19.6	1179	2	A53213
7	1102.5	18.8	1151	2	A45226
8	1084	18.5	1170	2	I45914
9	1072	18.2	1178	2	S44142
10	1069	18.2	1181	2	A33998
11	1060	18.0	1180	2	A35854
12	663	11.3	1039	2	A41131
13	633	10.8	1038	2	S06046
14	630	10.7	1035	2	I58409
15	614.5	10.5	1041	2	T31437
16	579.5	9.9	1054	2	JC7294
17	572.5	9.7	1051	2	A35761
18	567.5	9.7	1053	2	I55534
19	555.5	9.5	1053	2	S44250
20	543.5	9.3	1034	2	A36108
21	535	9.1	1044	2	T10050
22	532	9.1	1049	2	B24079
23	532	9.1	1073	2	B36429
24	530.5	9.0	1072	2	A38457
25	529.5	9.0	1051	2	A40021
26	526	9.0	1048	2	A27421
27	525.5	8.9	1091	2	A41543
28	517	8.8	1044	2	S16516
29	496.5	8.5	1146	2	S40311

ALIGNMENTS

RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567

J:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3,

B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:CROSS-references: GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148

A:Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499, 501-965, 'P', 967-1153 <ARN>

A:CROSS-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA5

A:Note: The authors translated the codon TAC for residue 1129 as Thr

A:Note: part of this sequence, including the amino end of the mature protein, was conf

R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally i

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:CROSS-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe

A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:CROSS-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept

A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:CROSS-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A:Reference number: A46526; MUID:93123748; PMID:8419480
A:Accession: A46526
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499,501-1153 <FLE>
A:CROSS-references: GB:S52227; NID:9263047; PIDN:AB24821.1; PID:9263049
A>Note: The last three bases of intron 13, CAG, are included in some but not all mature
A>Note: sequence extracted from NCBI backbone (NCBIP:121963)
R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A:Reference number: A90664; MUID:87076671; PMID:3539202
A:Accession: A26091
A:Molecule type: protein
A:Residues: 17-31 <PIE>
A:Experimental source: granulocytes
R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A>Title: Characterization of the myeloid-specific CD11b promoter.
A:Reference number: 152567; MUID:92144986; PMID:1346576
A:Accession: 152567
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-9 <RES>
A:CROSS-references: GB:M84477; NID:g180184; PIDN:AA51960.1; PID:g553219
C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C:Genetics:
A:Gene: GDB:ITGAM; CR3A
A:CROSS-references: GDB:120599; OMIM:120980
A:Map position: 16p11.2-16p11.2
A>Note: Promoter contains a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F:17-1108/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <VMA2>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.9% Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVAAQNRGLYQCDYSGCEPI 60
Db 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVAAQNRGLYQCDYSGCEPI 76

Qy 61 RLQDPVEAVNMSLGLSLAATSPOLLACGPTVHTCSENTYVVKGLCFPGSNLRQQPK 120
Db 77 RLQDPVEAVNMSLGLSLAATSPOLLACGPTVHTCSENTYVVKGLCFPGSNLRQQPK 136

Qy 121 PPEARLGCPOEDSDIAFLIDSGSIIPHDFRMKEFVSVVMEQLKSKTLFSLMOYSEEF 180
Db 137 PPEARLGCPOEDSDIAFLIDSGSIIPHDFRMKEFVSVVMEQLKSKTLFSLMOYSEEF 196

Qy 181 RIHFTFKFQNNPNRSLVKPTQLLGRTHATGVRKVIIRRELLNITNGARKNAFKILVI 240
Db 197 RIHFTFKFQNNPNRSLVKPTQLLGRTHATGVRKVIIRRELLNITNGARKNAFKILVI 256

Qy 241 TDGEKFGDPLGVEDYVPEADREGVIRYVIGVDAPFRSEKSKQELNTIASKPPRDHFQVN 300
Db 257 TDGEKFGDPLGVEDYVPEADREGVIRYVIGVDAPFRSEKSKQELNTIASKPPRDHFQVN 316

Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGPLLSTVGSYDWAG 376

Qy 361 GVFLYTSKESKSTFINNTRVDSNDMDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINNTRVDSNDMDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEOTRGQOVSCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEOTRGQOVSCPL 496

Qy 481 PRGQARWQCDAYLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDAYLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGOSLGGODLTMDGLVDLTVGAGQHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGOSLGGODLTMDGLVDLTVGAGQHVLLRSQ 616

Qy 601 PVLVRKAIMEFNPREVARNVFECDVVGKGEAGEVRVCLHVQKSTRDRREGIOISVVT 660
Db 617 PVLVRKAIMEFNPREVARNVFECDVVGKGEAGEVRVCLHVQKSTRDRREGIOISVVT 676

Qy 661 YDLALDSGRPHSAVFNETKSTRQTQVLGLTQTCETLKLQPLNCIEDPVSPIVRLNF 720
Db 677 YDLALDSGRPHSAVFNETKSTRQTQVLGLTQTCETLKLQPLNCIEDPVSPIVRLNF 736

Qy 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALPFPEKNCNDNICDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEDAQRFLTALPFPEKNCNDNICDDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKVKYSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKVKYSTLQNRQSRQSWRLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916

Qy 901 QLELPVKYAVYVMVTSHGVSCTKYNLFTASENTSRVMOHQYQVSNLQSRSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVMVTSHGVSCTKYNLFTASENTSRVMOHQYQVSNLQSRSLPISLVFLVPV 976

Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036

Qy 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSQTET 1080
Db 1037 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSQTET 1096

Qy 1081 KVPEFVFNPLPLIVGSSVGLLALLITALYKLGFFKRYKQKDMWSEGGPPGAEPO 1137
Db 1097 KVPEFVFNPLPLIVGSSVGLLALLITALYKLGFFKRYKQKDMWSEGGPPGAEPO 1153

RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N:Alternate names: complement-3 receptor alpha chain

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 22-Oct-1999

C:Accession: S00551; I59078

R:Pytela, R.

EMBO J. 7, 1371-1378, 1988

A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with th

A:Reference number: S00551; MUID:88312584; PMID:3044779

A:Accession: S00551

A:Molecule type: DNA

A:Residues: 1-1153 <PYT>

A:Cross-references: EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52983

A>Note: the authors translated the codon CAC for residue 569 as Gln

R.Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
A:Reference number: I59078; MUID:86287312; PMID:2942940
A:Accession: I59078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.2%; Score 4476; DB 2; Length 1153;
Best Local Similarity 74.3%; Pred. No. 1.2e-301;
Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;

QY 1 FNLDTENANTFOENARFGQSVVVOLOGSRVVVCAQOEIVAAVNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTEHPMTFOENAKFGQSVVVOLOGSRVVVCAQOEIVAAVNRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATSPQALLACGPTVHOTCSNTVYVKGFLCFLFGNLRPOOK 120
DB 77 PLQVPPEAVNMSLGLSLAVSTVPQALLACGPTVHQNKENTVYVKGFLCFLFGNLRPPQ 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFEVSTVMEQKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIVFLIDGSGSINNIDFQMKFEVSTVMEQKSKTLFSLMQYSEEF 196

QY 181 RHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIRELLNITNGARKNAFKLLIYI 240
DB 197 RHFTFNDPKRNPSPRSHVSPKQLNGRTKTASGIRKVRRELPHKTNAGRENAAKILYI 256

QY 241 TGEKFGDPLGVEDVTPEDAREGVIRYVIGVGDAPRSEKROBLNTIASKPRDHVQV 300
DB 257 TGEKFGDPLGVEDVTPEDAREGVIRYVIGVGDAPRSEKROBLNTIASKPRDHVQV 316

QY 301 NFEALXTIQNLEKIFAJEGTQTGSSSFEHMSQEGFSAAITSNGPLSTVGSYDMAG 360
DB 317 NFEALXTIQNLEKIFAJEGTQTGSSSFEHMSQEGFSAAITSNGPLSTVGSYDMAG 376

QY 361 GVFLYTSKSKSTFINNTRVDSNDMDAYLGVAAILRNVRQSLVILGAPRYQHIGLVMPR 420
DB 377 GAFLYTSKSKSTFINNTRVDSNDMDAYLGVAAILRNVRQSLVILGAPRYQHIGLVMPR 436

QY 421 QNTGMSEANVKGTOIGAYFGASLCSVDVDSNGSDLVILGAPHYVEOTRGQSVSCPL 480
DB 437 ENFGTWEPTSTIKGSQLGSGVFGASLCSVDVDSNGSDLVILGAPHYVEOTRGQSVSCPL 496

QY 481 PRQARWQCDALVYGEQGPWCRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAVILF 540
DB 497 PRG-RARWQCDALVYGEQGPWCRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAVILF 555

QY 541 HGTSGSGISPHSHQRTAGSKLSPRLQYFGQSLSGGQDLTDGLVLDITVGAQGHVLLRSQ 600
DB 556 YGASIASLASASHSHRIIGAFHSPGLQYFGQSLSGGQDLTDGLVLDITVGAQGHVLLRSQ 615

QY 601 PVLRVKAIMEFNPREVARNVFECDVWVGKEAGEVRVCLVHVKSTRDLREGQIOSVVT 660
DB 616 PVLRVKAIMEFNPREVARNVFECDVWVGKEAGEVRVCLVHVKSTRDLREGQIOSVVT 675

QY 661 YDLALDSGRPHSAFVNETKNSRRQTQVLGLTQTCTETKLOLPNCIEPVPSPVILRNLF 720
DB 676 YDLALDSGRPHSAFVNETKNSRRQTQVLGLTQTCTETKLOLPNCIEPVPSPVILRNLF 735

QY 721 SLVGTPLSAFNGLRPVLAEDAQRLLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 736 TLVGEPLRSFGNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 795

QY 781 GPREFNVTVVRNDEGSDYRQVTFPPFLDLSYRKVSTLQNRQSRL-ACESASSTE 839
DB 796 GPQDNMSVTLNRNDEGSDYRQVTFPPFLDLSYRKVSTLQNRQSRL-ACESASSTE 855

QY 840 VSGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNTE 899
DB 856 GHGALKSTTWNINHPFPANSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNTE 915

QY 900 FQLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVFLVP 959
DB 916 FQLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVFLVP 975

QY 960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVQORIQCDI 1019
DB 976 VQINNVTWDRHPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVQORIQCDI 1035

QY 1020 PFFGLOEENFATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOTE 1079
DB 1036 PSFNTQEIFNVTLKGNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOTE 1095

QY 1080 TKVEPFEVNPVPLITVSGSVGGLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1137
DB 1096 TKVEPFEVNPVPLITVSGSVGGLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1153

RESULT 3
RWHTUC
cell surface glycoprotein CD11c precursor - human
N:Alternate names: leukocyte adhesion receptor p150.95 alpha chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999
C:Accession: A35584; A35543; S00864
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A35584
A:Contents: erratum
A:Accession: A35584
A:Molecule type: DNA
A:Residues: 1-1163 <COR>
A:Note: this revision to the sequence from reference A35543 includes the carboxyl end
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150.95 molecule
A:Reference number: A35543; MUID:90153906; PMID:2303426
A:Accession: A35543
A:Molecule type: DNA
A:Residues: 1-834 <CO2>
A:Note: this sequence has been revised in reference A36584
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A:Reference number: S00864; MUID:88166645; PMID:3327687
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755, 'L', 757-1163 <CO3>
A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119758; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F:120-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status F

Query Match	59.3%;	Score 3483;	DB 1;	Length 1163;
Best Local Similarity	61.4%;	Pred. No. 8.8e-233;		
Matches	693;	Conservative 138;	Mismatches 292;	Indels 6; Gaps 4;
Qy	1	FNLDTENAMTFQENARGFGOSVVVLOGSRVVVGAPOEIVAANQORSLYQCDYSTGSCRPI	60	
Db	20	FNLDTBELTAFRVDSAGFGSDVVQYANSWVVGAPQKITAAQTGGLYQCGYSTGACEPI	79	
Qy	61	RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQPOK	120	
Db	80	GLQVPPVAVNMSLGLSLASTSPSQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTOR	137	
Qy	121	FPEALRCQOEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMQOLKSKTLFSLMQVSEEF	180	
Db	138	LPVSRQECPRQEQDILVFLIDGSGSISRNFATMNFVRAVISQFQRPSTQFSLMQVSNKF	197	
Qy	181	RIHFTFKFQNNPNRSLRKPITQLLAGRTHATGVRKVIRELLNITNGARKNAFKLIVI	240	
Db	198	QTHFTFEFRRTSNPLSLASVHQLQGFYTATAIQNVHRLPHASYGARRDATKILIVI	257	
Qy	241	TGDEKFGDPLGYEDVIEADREGVIRYVIGVDAPFSEKSRQELNTIASKPPRDHVPQVN	300	
Db	258	TGDKKSGDSLKYKDVIMADAAIIRYAGVGLAFQNRNSWKLNDIASKPSQEHIFKVE	317	
Qy	301	NFEALKTIQNLREKIPAIETGOTGSSSFEHMSQEGFSAATISNGPLSTTVSGYDWAG	360	
Db	318	DFDALKDIONLKEKIFAIEGTETSSSFELEMAQEGFSAVFTPDGFLVCAVGSFTWSG	377	
Qy	361	GVFLYTSKEXSTFINMTRVSDMDNDAYLGAAAAIILNRVQSLVGLCAPRYOHIGLVAMFR	420	
Db	378	GAFLYPPNMSPTFINMSQENVMDSDYLGSTELALWKGVQSLVGLCAPRYOHTGKAVIFT	437	
Qy	421	QNTQWESNANVKGTQIATFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGGQVSCPL	480	
Db	438	QVSRQWRMAEVTGTQISYFGASLCSVDVDTGSDTLVLIGAPHYYEOTRGGQVSCPL	497	
Qy	481	PRGORARWQDAVLYGQGGQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEENNRGAYILF	540	
Db	498	PRGWR--RWQDAVLYGQGGQWGRFGAALTVLGVDVNGDKLTDVVGAPGEENNRGAYILF	556	
Qy	541	HGTSGSGISPSHSORIASGLSRLOVFGQSLSGGQDLTMDGLAVLTVGAQGHVLLRSQ	600	
Db	557	HGVLGPSISPSHSORIASGLSRLOVFGQSLSGGQDLTQDGLVLDVAVGARGQVLLRTR	616	
Qy	601	PVLRVKAIMENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRREGIQIQSVVT	660	
Db	617	PVLVWGVSMQFIPIAEIPRSFAFEQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT	676	
Qy	661	YDLALDSRPHSRVAFNETKNSRROTQVGLGLTQTCETLKLQLENCIEDPVSPIVLRNLF	720	
Db	677	LDLALDPGLRSLPRATFQETKNSLSRVRLVGLKAHCENFNLLPSCVEDSVPTILRLNF	736	
Qy	721	SLVGTPLSFAFNLRPLVLAEDAQRLLFTALFPFEKNCNGNDNIQDDLSITFSFMSLDCLVWG	780	
Db	737	TLVGKPLAFNLRPLVLAEDAQRVFTASLPEKNCNGADHIQDNLGISFSFPGKSLLVG	796	
Qy	781	GPREFNVTVVRNCGEDSYRTQVTFPPDLISYKRVSTLQNRQSRWRLACESASSTEV	840	
Db	797	SNLELNAEVMVWMDGEDSYGTTITFSHPAGLSYRVYAEQGGQQLRSLHLTCDSPAGV--	854	
Qy	841	SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTFF	900	
Db	855	SGQTWSTSCRINHPIFRGGAQITFLATFDVSPRAVLGRDLTLTANVSSENTPRTSKTTF	914	
Qy	901	QLELPVKYAVVMVTVSHGVSTKYLNFTAS--ENTSRVMQHQVQSNLQORSPLISLVFLVP	959	
Db	915	QLELPVKYAVVTVVSSHEQTKYLVNFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP	974	
Qy	960	VLRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVNVCISIAVQRIQCDI	1019	
Db	975	VELNQEA VMVDEVSHPQNSLRCSSEKIAAPPASDFLAHQVPLVDCSIAGCLURFCDV	1034	
Qy	1020	PFQIQBEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPGQAFVRSQTE	1079	

Db	1035	PSFSVQEEELDTLKGNSLFGWVRQLQKKVSVVVAEITFDTSVSQLPQGBAFMRAQTT	1094		
Qy	1080	TKVEFEVFNPLPLIVGSSVGLLILALITAAALYKLGFEKROYKDMWSE	1128		
Db	1095	TVLEKYKVHNPTPLIVGSSVIGGLLLALITAVLYKVGFPKROYKEMWEE	1143		
RESULT 4					
S03308					
cell surface glycoprotein CD11a precursor - human					
C;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function-associated molecule-1 alpha chain					
C;Species: Homo sapiens (man)					
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 20-Aug-1999					
C;Accession: S03308; A47458; A48759; S36044					
C;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.					
J. Cell Biol. 108, 703-712, 1989					
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha chain					
A;Reference number: S03308; MUID:89139587; PMID:2537322					
A;Accession: S03308					
A;Molecule type: mRNA					
A;Residues: 1-1170 <LAR>					
A;Cross-references: EMBL:Y00796; NID:931421; PIDN:CAA68747.1; PID:931422					
A;Note: part of this sequence was confirmed by protein sequencing					
R;Cornwell, R.D.; Gollahon, K.A.; Hickstein, D.D.					
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993					
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a)					
A;Reference number: A47458; MUID:93248261; PMID:8097887					
A;Accession: A47458					
A;Molecule type: DNA					
A;Residues: 1-20 <COR>					
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)					
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.					
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993					
A;Title: Identification of cell-specific and developmentally regulated nuclear factor					
A;Reference number: A47565; MUID:93281759; PMID:8099450					
A;Accession: A47565					
A;Molecule type: DNA					
A;Residues: 1-20 <SHE>					
A;Cross-references: GB:M95609					
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.					
J. Biol. Chem. 268, 19305-19311, 1993					
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter					
A;Reference number: A48759; MUID:93374910; PMID:8103515					
A;Accession: A48759					
A;Molecule type: DNA					
A;Residues: 1-20 <NUE>					
A;Cross-references: EMBL:Z22804; NID:9311405; PIDN:CAA80461.1; PID:9311406					
C;Genetics:					
A;Gene: GDB:ITGAL; CD11A					
A;Cross-references: GDB:119757; OMIM:153370					
A;Map position: 16p11.2-16p11.2					
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat					
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen					
F;1-25/Domain: signal sequence #status predicted <SIG>					
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted					
F;154-317/Domain: von Willebrand factor type A repeat homology <VMA2>					
Query Match 26.4%; Score 1548.5; DB 2; Length 1170;					
Best Local Similarity 34.5%; Pred. No. 1.2e-98;					
Matches 405; Conservative 208; Mismatches 460; Indels 101; Gaps 37;					
Qy	1	FNLDTENAMTFQ--ENARGFGSVVQLOGSRVVVGAPOEIVAANQSGSLYQCDYSTGSCSE	58		
Db	26	YNLDRVGRKSFSPPRAGRHFYRVLQV--GNGVTVGAPGE---GNSTGSLYQCSGTGCHCL	81		
Qy	59	PIRLQVPVAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFPGSNLR---	115		
Db	82	PVTLR--GSNTYSKYLGWTLATDPTDGSILACDPLGSLRTCDQNTYLSGLCYLFRQNLQCPM	140		
Qy	116	-QOPQKPEALRCPOEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQKSKTILFSLM	174		
Db	141	LQGRPGFQECIKG---NVDLVFLFDGSMSLQDPDEFQKILDFMKDVMKKLSNTSYQFAAV	196		

QY 763 DLSITSPMSLDCLVVGGR-----REFNVTVTRNDGSDSYRTQVTFPPPLDLSYRKV 816
D 775 ANLTLSPPARS-----GRLMSSASLAVETWLSNGSDAYWRDLDFPRGLSPRKV 827
QY 817 STLQNSQSRWRLACASSTEVSGAL-KSTSCSINHPIPPENSEVTFNITFDVDSKAS 875
D 828 EMLQ---PHSRMPVSCBEL--TEGSSLLTKLKNVSSPIFKAGQEVSLQWMENTLNS 882
QY 876 LGNKLKLLKANTSEN-NMPRTNKTEFQLELPVKYAVVMVTSKYNLFTASENISR 934
D 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLVPRNLTKQENSTLYISFPGPKQTQ 942
QY 935 VMQHOYOVNLSQSLPISLVLVPRNLQVIMDRFO---VTFSENLS---TCHTK 986
D 943 QVQHYQV---RIQSAVDHNMPT-LEALGVPRPHSEDLITYTSVQTDPLVTCHE 996
QY 987 E-RLPSSHDFLAELRKAPVWNCISIAVCQIQDIPFFGIOBEFNATLKGNSLDFWYIKTS 1045
D 997 DLKRPSE--AEQCLPGV-----QFRCPVIF---RWEILIQVTGTVELSKETKAS 1042
QY 1046 HNHLIIVSTAILFNDSVFTLLPGCAFVRSQTEKTEKPEVPPNPPLPLIVGSSVGGGLLL 1105
D 1043 -STLSLSSLSVSNSSKHFLYGSKA-SEAQVLVKVDLIHEKEMLHYVLSGGGLVLL 1100
QY 1106 ALITAALYKLGFEKRYQKDM-SEGGPPGAP 1136
D 1101 FLIFALYKVGFFKRNLEKMEADGGVNGSP 1132

RESULT 6
A53213
integrin alpha-5 chain - human
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: A53213
R:Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (E) subunit. Ur
A:Reference number: A53213; MUID:94164962; PMID:8119947
A:Accession: A53213
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: GB:L25851; MID:9457244; PID:9457245
C:Genetics:
A:Gene: GDB:ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homod
F:199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.6%; Score 1149; DB 2; Length 1179;
Best Local Similarity 29.1%; Pred. No. 5.9e-71;
Matches 343; Conservative 211; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSLYQDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPPOLLACGPTVHQ 95
D 65 GLPRLCSLVQDEILCHPVEHVPDKGRHRGVTVVRSHHGVLCI-----QVLVRRP--HS 117
QY 96 TCSENTYVKGILCFGLSGLRQPOQ-----LLENDLPDARDVDTGDCYSKKEGGEDDVNTA 174
D 118 LSSELT---GTCSLGLPDLRPOAQANFDFLENLDPDARDVDTGDCYSKKEGGEDDVNTA 174
QY 120 KPEALRGCPQED-----SDIAFLIDGSGIIPHDPRMKEFVSTVNEQL-- 164
D 175 RRRALKEEEDKEEEDDEEAEAGTEAIIIDGSGIDPPDFORAKDFISNMNRNFE 234
QY 165 KSKTLFSLMQYSEFRHTFKEQNPNPRLVKPIITOLLGRHTATGVRKVIKRELLN 224
D 235 KCFECNFALVQGVGIQTEFDRDSQDVNASLARVONITQVSGVTKTASAMOHVLDISFT 294
QY 225 ITNGARKNAFKILIVITDCEKFGDPLGVEDVPEADREGVIRVYGVGDADFSEKSRQEL 284

RESULT 7

A45226

integrin alpha-1 chain - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 31-Dec-2000

C:Accession: A45226

R:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.

J. Biol. Chem. 268, 2989-2996, 1993

A:Title: Expression of native and truncated forms of the human integrin alpha 1 subur

A:Reference number: A45226; MUID:93155124; PMID:8428973

295 SSHGRRKASKVMVVLTDGIFEDPLNLTTVINSKMQQGVRFPAIGVGEEFKSARTAREL 354
QY 285 NTIASPPRRDHVFOVNNPEALKTIONOLREKIPAEIGTOTGSSSSSEHEMSQEGFSAAIT 344
D 355 NLIASDPDTHAFKVTNTYHALDGLLKLRYNIISHEGT---VGDALHYQLAQIGFSAQIL 411
QY 345 SNGP-LLSTVGSDYDAGGVFLY-TSKEKSTFINMTRVSDMDND-----YLGAAAIILRN 398
D 412 DERQVLLGAVGAFDMSGGALLYDTRSRGRFLNQTAATAAADAEEAAQYSLGVAVAVLHKT 471
QY 399 RVQSLVLGAPRVQHIGLVAMFR-ONTGMHESNANV-KGTQIGAYFCASLCSVDVDSNGST 456
D 472 CSLSYVAGAPQYKHG--AVFELQKEGRASFLPVLEGEOMGSGYFSELCPCVDIDMDGST 529
QY 457 DLVLICAPHYEQTRCGQVSVCPPLPRGARWQCDVAVLGEQCPMGRFCAALTVLGDVN 516
D 530 DFLVVAAPFYHVHGEGRVTVYRLSE-QDGSFSLARILSGHPQPTNARFGFAAAMGDL 588
QY 517 GDKLTDAVAGAP---GEEDNR--GAVLYPHGTSGSGISPSHSQRIAGSKLSPLRYFGQ 570
D 589 QDKLTDAVAGAPLEGFGADDGASFGSVIYNG-HWDGLSASPSQIRASTVAPGLQYFGM 647
QY 571 SLSSGQDLTMDGLVDLTVGAQGHVLLLRGQVPLRVKAIMFENPREVARNVFECNDQVKG 630
D 648 SMAGGFDISGDGLADITVGTGQAVVFRSRPVVRLKVSMAFTPSALP-----IGF 697
QY 631 KEAGEVVRVCLHVQKSTRDLREQIOISVVTVYDIALDSGRPHSRVAFNETKNSTRRTQVL 690
D 698 NGVNVRLCFEI--SSVTASESLREALNFTLDVVGKORRLQDVSRCGLCREWS 756
QY 691 GLTQTCETLKLQLPN---CIEDPVGPVILRLNFSLVGTPLSAFGLNLRPLVLAEDARLFT 746
D 757 SGSQLCEDL-LLMPTGELCEEDCFNSASVKVSYQL-QTPEGQTDHPQPLDRLRYTEPFAI 814
QY 747 ALPPEKNCNDNI CODDLSITFSPSLDCLVVGGRREFNVTVTRNDGSDSYRTQVTF 806
D 815 FQLPYEKACKNKLFCVAELQLA-TTVSQQELVVLGTLKELTLNLTNSGSDSYMTSMALN 873
QY 807 FPLDLSYRKVSTLQNSQSRWRLACASSTEVSGALKSTCSINHPIPPENSEVTFNI 866
D 874 YPRNLQ-----LKRQKPPSPNIQDDPQPV---ASVLNMCRIQHPVL-KRSSAHVSV 923
QY 867 TFDVDSKASLGNKLLKANVTSEN---MPTNKTEFQ---LELPVKYAVYVMVTSHG 919
D 924 VQOLEENAFNRTADITVTVTNSNRRSLANETHTLQFRHGFVAVLSKPSIMYVNTCQGL 983
QY 920 S---TKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPLVRLNQTIVWDRPQVTFSE 977
D 984 SHHKEFLFHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAAVKLRTQ 1028
QY 978 NLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQIQDIPFFGIOBEFNATLKGNSL 1037
D 1029 ASTVCTWSQBRACAYSS-VQHVEMHSHVSCVIA-----SDKENVTVAEIS 1073
QY 1038 FDMYIKTSHNHLIVST-----AEILFNDSVFTLLPGQAGFVRSQTEKTEKPEVPEVNP 1091
D 1074 WD-----HSEELKQVTELQILGEISFNKSLVEGLNAENH--RTKITVTVFLKDEKYHSL 1125
QY 1092 PLIVGSSVGGLLLALITAAALYKLGFEKRYQKDMSE 1128
D 1126 PIITKGVGGLLVILVILVILFKCGFFKRYQQLNLE 1162

A;Accession: A45226

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1151

A;Experimental source: hepatoblastoma cell line HepG2

A;Note: Sequence extracted from NCBI backbone (NCBIP:124326)

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F;142-317/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 18.8%; Score 1102.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 9.4e-68; Indels 195; Gaps 44;
Matches 344; Conservative 211; Mismatches 487;

QY 1 FNLDTENAMTFOENARG-FQSGVVQL---QSGRVVVGAPQEIIVAAANQORSLVQCDYSTGS 56
DB 1 FNVVKNMSTFSPVEDMEGYTVQYENEBGKVLIGSLVGPQKRTGDVYKCPVGRGE 60
QY 57 CEP-IRLVQVVEA-----VNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVKG 106
DB 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGGFLACGPLYAYRCGLHLHYTTGI 119
QY 107 CFLFGSNLRQOPKFPEALRGCPQEDSDTAFLIDGSGSIIPHDPRMKFEFVSTVMEQLK- 165
DB 120 CSDVSPTFQVNVSIAP-VQECSTQ-LDIVILDGNSIYPWD--SVTAPLNDLLKRMDDI 174
QY 166 -KSKTLFLSMOYSEEFRIHFTFKFQNNPNRSLVKPITQLLGR-THTATGVRKVIKRELL 223
DB 175 GPKQTQGVIOYGENWTHFENLNKYSSTEVLVAAKIVORGROQTWALTGDTARKEAF 234
QY 224 NITNGARKNAFKILIVITDGEKFGPLGYEDVPEADREGVIRYVIGVGDAFR-----SE 278
DB 235 TEARGARRGVKVMVITDGEH-DNHRLKKVIOCDENIQRFSAIILGSYVNRGNLSTE 293
QY 279 KSRQELNTASKPRPHVQVNNFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQSG 338
DB 294 KFVEEKISASEPTEKHFNVSDALVTIVKTILGERIFALEATADQSAASFEMMSQTG 353
QY 339 FSAATISNGPLLTGVSYNWAGVFLVTSKE-----KSTP-INMTRVDSMDMDAYLGAA 392
DB 354 FSAHYSDWMLGAVAYDWNGVVMQKASQIIIPNTTFNVSTKKNPL-ASYLGITYV 412
QY 393 AITLRNRVOSL-VLGAPEYQHIGLVAMFRONTGMESNANVKGTQIGAYFGASLCSVDVD 451
DB 413 NSATASSGDVLYIAGQPRYNTQGVIIYRMEDGNIKILOTLSEQIGSYFGSLITTTDID 472
QY 452 SNGSTDLVLIGAPHY-----YEQTR-GGQSVSCPLPRGORARWQCDVAVLY 495
DB 473 KDSNTDILLVGAPEYMGTEKEEGQKVYVVALNQTREYQMSLEPIKQTCSSRQHNSCTT 532
QY 496 GEQGPWG-RFGAALTVLGDVNGDKLTDVAIGAAGEDNRGAVYLFHGTSGSGISPSHSQ 554
DB 533 ENKNEPCGARFGTAIAVNDLNDGNDIVIGAPLSDDHGGAVYIYHG-SGKTIKRYAQ 591
QY 555 RIAGSKLSPRLQYFGOSLSGGDLTMDGLVDLTGVAQGHVLLLRSPVLRVKAIMEFNPR 614
DB 592 RIFSGDGDKTLKFFGSGIHGEMDLNGDGLTDTVIGLGGALPWSRDAVVKVWTFEPN 651
QY 615 EVARNVPCNDQVVKGEAG-EVRVCLHVQ-KSTRDRLEGQIQSVVYVDLALDSGRPH 671
DB 652 KVNIOKKNCH---MEKETVCINATVCFVKLSKEDTIEADLQ---YRVTLDSLRQI 704
QY 672 SRAVENET-----KNSTRQTOVLGTQTCETLKLQLPNCIEDPVPVILRLNFSLVGT 725
DB 705 SRSFSGTQBRKVQRMNITVRKSEC-----TKHSFYMLDRHDFQDSVR---ITUDFNLT-D 755
QY 726 PLISAFGNLRPLVAEDAQLFTALFPPEKNCNDNIQODLSIIFSPMSLDCLVVGPRE- 784
DB 756 PENG-----PVLDSLPNSVHEIYPAKDCGNKEKICISLDSLVHATTEKDLLIVRQNDK 810
QY 785 FNVTVVRNDGDSYRTQVTFPFLDLSVRKVKSTLQNRQSRQSRWLACESASSTEVSGAL 844
DB 811 FNVSLVKNTKDSAYNTRTIVHVSPLNVFSGIEAIQKD-----SCSEN-----853

QY 845 KSTSCSINHPIPPENSEVTFNITFDVDSKASIGN-KLLLLKANVTSENMPRTNKTFFOLE 903
DB 854 HMITCKGVYFLRRGEMVTFKILFQNTSYLMENVTIVLSATSDEEPPETLSDNVNVIS 913
QY 904 LPVKYAVMVVTSHGVSFKYLNFTASENTSRVMQHOYQVSN-----LQORS-----L 950
DB 914 IPVKYEVGLQFYS-SASEYHISIAAANTVPEVINSTEIDIGNEINIFYLIRKSGSPMPBL 972
QY 951 PISLVF-----LVPVRLNQTVIWDPRQVTFSENLSSTCHTKE-----RLPS 991
DB 973 KLSISFPNMTSNGYVPVLYPTGLSS-----SENACRPHIFEDPFSINSKGMTT 1021
QY 992 HSDFLAEIRKAPVNCVSIACQRIQCDIPFFGQIE-----EFNATLK 1033
DB 1022 STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILWKPTFKYSFSSNLNLTIR 1078
QY 1034 GNLSDFWIKTSHNHLILVSTAEIILPNDVSFTLLPGQAGFVRSQETETKVEPEVNPPL 1093
DB 1079 GEL-----RSENASLVSSN-----OKRELAIOISKDGLPGRVPL 1114
QY 1094 --IVGSSVGGLLLLALITAAALYKLGFFKQYKDMSE 1128
DB 1115 WVILLSAPAGLLLLMLLILALWKIGFFRPLKKKMEK 1151

RESULT 8

I45914

Integrin alpha 2 subunit - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Sep-1999

C;Accession: I45914

R;Kamata, T.; Puzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A;Title: Identification of putative ligand binding sites within the I-domain of integrin

A;Reference number: A54402; PMID:94193847; PMID:7511592

A;Accession: I45914

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1170 <KAM>

A;Cross-references: GB:L25886; NID:9439695; PID:AA859255.1; PID:9439696

C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology

F;161-336/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.5%; Score 1084; DB 2; Length 1170;
Best Local Similarity 27.7%; Pred. No. 1.9e-66;
Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFO-ENARGFQSGVVQL---QSGRVVVGAPQEIIVAAANQORSLVQCDYST 54
DB 19 YNVGLPKAKIFSGPSSEQGYAVQQFINPKGNWLLVGSWSPGPKNRMGDVYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVKG 107
DB 79 TTCEKLNLTQSTSMNSVTEMTKNSLGLTLTRNVGTGGFLTCGPLMAOQSGSYTTGVC 138
QY 108 FLFGSNLRQOPKFPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKFEFVSTVMEQLK-- 165
DB 139 SDVSPDF-QLRTSPAPAVQTCF-SFIDVVVCDSENSIYPWD--AVKNLEKVFQGLDGLG 194
QY 166 KSKTLFLSMOYSEEFRIHFTFKFQNNPNRSLVKPITQLL-----GRTHATGVRKVIKRE 221
DB 195 PKTKQMLLIQYANNPRVFNLTNFKSKD---EMIKATSTQTFQYGGDLTNTFKAIOVARDT 251
QY 222 LNIINGARKNAFKILIVITDGEKFGPLGYEDVPEADREGVIRYVIGV-----GDAFR 276
DB 252 AYSTAAGRPGATKVMVVVTDGESH-DGSKLXAVIDQCNKNILRFGIYGLVGLYNNALD 310
QY 277 SKSRQELNTIASKPRDRHVFOVNNFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQ 336
DB 311 TKNLKIKAIASIPTEHFHFNVSDEADLEKAGTIGEIQFSLTEGVQO-GNFFQNMESQ 369
QY 337 EGFSAAIT--SNGPLISTVGSYDWAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGVA 391

Db 370 VGFSAEYSPQNNILMLGAVGAYDMSGTVVQKTPHCHLIFSKQAEQIQLDRNHSSYLGS 429
Qy 392 AAILLRNVQSLVGLGARYQHIGLVAMFRQNTGWMESNANV-----KGTQIGAYFGASL 445
Db 430 VASISTGNSVHFVAGAPRANTGTQIVLYSVN-----ENGNNVTIQSORQGIYFQSVL 484
Qy 446 CSVDVDSNGSDLVILGAPHYIEQTR--GGQVSVCPPLRGORARWQCDVLYEGQGPWG 503
Db 485 CAVDNKTITDVLVGLGAPMYMNDLKKEBGRVYLFTITKG-ILNWH--QFLEGPNGLNA 541
Qy 504 RFGAALTVLGVNGDKULTDVAIGAPGEDNKGAVYLFHGTSGSISPSHSGRIAGS--KL 561
Db 542 RFGSAIAALSIDNMGDFNDVTVGSPLENQNSGAVIYNGHEGM-IRLRYSQKILGSDRAF 600
Qy 562 SPRLOYGQSILSGQDLTMDGLVLTGCAQGHVLLRSQVLRVKATMEFNPREVARNVF 621
Db 601 SSHLOYGERSLDGVDGLNGDSITDVSAGAFQVQVQLWSQSADVSDASTFPKRI--TL 658
Qy 622 ECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVTVYDLALD----SGRPHSRAVEN 677
Db 659 NKNAEI-----KLKLCF-----SAKFRPTNNQVAVIYNTIDEDQFSSRVSRLGFK 707
Qy 678 ETKNSTRQTVGLGTQTCB--TLKLQLPNCIEDPVPSPVLRNLFSL--VGTPLSAPGNL 733
Db 708 ENNERCLOKTMIVSQACRCSEYIIHIOEPS---DIISPLNLCMNISLENPGT----- 756
Qy 734 RPLVAEDAQRLFTALFPKCKNGNDNICODDLSITF---SFMSLDCLVGGPREFNVTV 789
Db 757 NPALAEVSETVKVFSIPPHKDCGDDGVCSIDLNVNQQLPATQQOOPFIVSNQKRLTFVS 816
Qy 790 TVRNDGDSYRTQVTFPPFLDLSYRKVSTLQORSQSRWLACESASST-EVSGALKSTS 848
Db 817 QLNKKSAYNTEIVVDESENLF-----ASWSPVDGTETVTCQIASQSKSVT 864
Qy 849 CSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRNTKTEFQLELPVKY 908
Db 865 CNVGYPAKSKQVTFITNFDNLQ-ILNQOASISFRALSQBENMADNSVNLKLSLLY 923
Qy 909 AVYVVTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQOR-----SLPISLFLV 958
Db 934 DAETHIT-RSTNINPYEVLGDNVSSV-HSFE--DIGPKFIPSKVTGTSVVPVSM----- 976
Qy 959 PVRLNQTIVDRPQVTFSEN---LSSTCHTK-----RSPSHSDFLAE- 998
Db 977 -----SVIHIPIQTKDNPLMYLTGVHTQAGDISCEAEINPLKIGQTSSSVFSKEN 1030
Qy 999 LRKAPVNCSTAVCQRIQCDIPFGIQBEFNATLKNLSFDWYIKTSHNHLILVSTAEI- 1057
Db 1031 FRHIKELNCRATASCNIMCWLRLDQVKGVEFLNVSTRIWNGTFAASTFQTVLQATAAEID 1090
Qy 1058 LPNDSVFTL-----LPQGAFAVRQSTETKVEPEE-VPNPLPLVIGSSVGLLALLLITA 1110
Db 1091 TYNPQIVYIENTVVTIP-----LTIMKPKHEVPTGVIGSVIAGILLALLLVA 1140
Qy 1111 ALYKLGFFKRYQVKOM 1125
Db 1141 ILMKLGFFKRYKYM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S44142

R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze

submitted to the EMBL Data Library, January 1994

A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not

A:Reference number: S44142

A:Accession: S44142

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <EDE>

A;Cross-references: EMBL:229987; NID:G473098; PIDN:CAA82877.1; PID:G473099
C;Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
P;169-344/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.2%; Score 1072; DB 2; Length 1178;

Best Local Similarity 28.1%; Pred. No. 1.3e-65;

Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;

Qy 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVVGAPOEIVAAHQRSLYOC--DYST 54
Db 27 YNVGLPKAKIFGSPSEQFGYSVQQLTNQGNWLLVGSFWSGPFENRMDGVYKCPVDLPT 86
Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSAAATTSPPOLLACGPTVHQTCSNTYVKGLC 107
Db 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLTCPLWAHQCGNYYATGIC 146
Qy 108 FLFGSNLRQOPO---KFPEALRCPCQEDSDIAFLIDGSGSIIPHDFRMKEFVSTWMEQL 164
Db 147 ----SDVSPDFQFLTFSFAVQACPSL-VDVVVVCDENSIYP--WEAVKNPLVKFVTGL 199
Qy 165 K--KSKTLFSLMOYSEEFRIHFTFKBFQNNPNRSLVKPITQLLG-RTHTATGVKRVIRE 221
Db 200 DIGPKKTOVALIOYANEPRILFNLDNFETKEDMVQATSETROHGGDLTNTFRAIEFARDY 259
Qy 222 LAMITNGARKNAFKILIVITDGEKFGDPLGYEDVPEADREGVIRYVIGV-----GDAFR 276
Db 260 AVSQTSGRPGATKVMVWVTDGESH-DGSKLKTVIQCCNDDEILRFGIAVLGLNRLNALD 318
Qy 277 SEKSRQELNTIASKPRDRHVFQVNNFEALQIONREKIFAIEGTQTGSSSSFEHMSQ 336
Db 319 TKNLKIKKAIKAIASPTPTERYFFNVADBAALLEKAGTIGEQIFSTEGTVQG--GDNFQWEMAQ 377
Qy 337 EGFSA--AITNSGPLLSTVGSYDAGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
Db 378 VGFSAADYAPQNDILMLCAVGAFDWSGTLVQETSHKPVIFPKQADFQVLDQRNHSSFLGYS 437
Qy 392 AAILLRNVQSLVGLGAPRYQHIGLVAMFRQNTGWMESNANV-----KGTQIGAYFGASLCS 447
Db 438 VAAISTEDGVHVFAGAPRANVTGQIVLYSVNK---QGNVTVIQSHRGDQIGSYFGSVLCS 494
Qy 448 VDVSNGSTDLVLIGAPHYIEQTR--GGQVSVCPPLRGORARWQCDVLYEGQGPWGRF 505
Db 495 VDVKDQITITDVLVIGAPTYNNDLKBEGKLYLTITKILNQHQ---FLEGPEGTGNARF 551
Qy 506 GAALTVLGVDVNGDKLTDVAIGAPGEDNKGAVYLFHGTSGSISPSHSGRIAGSKLSR- 564
Db 552 GSAIAALSIDNMGDFNDVTVGSPLENQNSGAVIYNGHQGT-IRTKYSQKILGSDNGAFRR 610
Qy 565 -LOYFGQSLSGGQDLTMDGLVLTGCAQGHVLLRSQVLRVKATMEFNPREVARNVFEC 623
Db 611 HLQFFGRSLDGYGDLNGDSITDVSIGALGQVQLWSQSIADVAIEALFTP----- 660
Qy 624 NDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVTVYDLALD----SGRPHSRAVENET 675
Db 661 -DKITLLNKDAKITLCLCFRAEFPRAGQNNQV--AILFNMTLDADGHSSVTSRGVFPREN 717
Qy 680 KNSTRROTQVLTQTCTET--LKLQLPNCIEDPVPSPVLRNLFSLVGTPLSAFGLNRPVL 737
Db 718 SERFLQKNMVVNEVQCKSEHHSIQKPS---DVNPLDLRVDLSLENPGTS-----PAL 768
Qy 738 AEDAQLFTALFPPEKCKNGNDNICODDLSI-----TFSPMSLDCLVGGPREFNVNT 789
Db 769 EAYSETVKVFSIPFYKEGSDGICISDLILDVQQLPAIQTSF-----IVSNQKRLTFS 823
Qy 789 VTVNDEGDSYRQVTFPPFLDLSYRKVSTLQORSQSRWLACESASST-EVSGALKST 847
Db 824 VILKNRGEYNTVTLAEFSENLF-----ASFSPMDVGTTEVTCEVSSQKSV 871
Qy 848 SCSTNHPIPFENSEVTENITFDVDSKASLGNKLLKANVTSENMPRNTKTE--FQLELP 905
Db 872 TCDVGYPAKSKQVQVTFITNFDNLQ-ILNQOAAINFQAFSESO--ETNKADNSVSLTIP 928
Qy 906 VKYAVVMVVTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQOR-----SLPISLV 955

```

Db 929 LLYDAELHLT-RSTNINFEISSDENAPSVIR---SVEDIGPKFIFSLKVTAGSAPVSN 984
QY 956 FLV-----PVLNQTIVMDRPOVTF--SENLS 980
Db 985 LVTHIPQYTKENPLLYLTGTQDQAGDISCTAEINPLKLPHTA-----PSVSPKVENPR 1040
QY 981 STCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIPFFGIQDEFNATLKGNSLFDW 1040
Db 1041 ---HTKE-----LDCRTTSCNITCMKDLHMKAEYFINTVTRVWNR 1080
QY 1041 YIKTSHNHLIVSTAEILLNDVFTLLPQGFAPVRSQETKVEPPEVENPLPLVGVSSVG 1100
Db 1081 FFASTFTQVQLTAADIDHNPOLFVIEENAVTIPLIMKPTKAEVPT--GVIIIGSIIA 1138
QY 1101 GILLALLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLALLAMTAGLWKLGFKQYKDM 1163

RESULT 10
A33998
integrin alpha-2 chain precursor - human
N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C:Species: Homo sapiens (man)
C>Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 21-Jul-2000
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha (2)-subunit (platelet
A:Reference number: A33998; PMID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <AK>
A:Cross-references: GB:X17033; NID:q33906; PIDN:CAA34894.1; PID:q33907
A>Note: the authors translated the codon GAT for residue 802 as Gin, GTC for residue 803
R:Catmell, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIC*, GPIIa an
A:Reference number: A56793; PMID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
A:Experimental source: platelet
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A:Reference number: A53117; PMID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, 'V', 18-21 <ZUT>
A:Cross-references: GB:L24121; NID:q400342; PIDN:AAA16619.2; PID:q4583535
A>Note: authors translated the codon GTA for residue 17 as Leu
C:Genetics:
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:128031; OMIM:192974
A:Map position: 5q11.1-5q11.2
A:Superfamily: unassigned collagens; von Willebrand factor type A repeat homology
C:Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CYT>
F:105,112,342,432,460,475,699,1037,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.2%; Score 1069; DB 2; Length 1181;
Best Local Similarity 27.0%; Pred. No. 2.1e-65;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43;

QY 1 FNLDTENAMTFQ-ENARGFGQSVQL----QGRVVVVGAPQETVAANQRGLYQC--DYST 54
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 YNVGLPEAKIFGSPSQEQGYAQVQFIPNPKGNLLWVSPGPPENRMGDVYKCPVDLST 89

```

```

QY 55 GSCPIRLQ-----VPVEAVNMNLSGLSLAATTPPOLLAGPFTVHQTCSNTYVKGLC 107
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 90 ATCEKLNLTSTSPNVTETKNTNLSGLILTRNMGTCGFLTCGPPLWACQCNQYVTTGVC 149
QY 108 FLFGSNLBOQKQKPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKFVSTVMEQLK-- 165
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 SDI9PDF--QLSASFPATQPCPSL--IDVVVVCDENSIYPWD--AVKNFLEKFGVGLDIG 205
QY 166 KSKTLFSLMQYSEBFRIFHTFEKFNNDNPRSLVKPITQLLG-RTHATATGVKVRRELLN 224
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 206 PTKQVGLIQLVANNPRVFNLTYYTKEEMIVATSTQSYGGDLTNTFGATQYARKAYS 265
QY 225 ITNGARKNAFKILIVITDGEKFGDPLGYEDVPEADREGVIRYVIGV-----GDAPRSEK 279
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 AASGGRRSATKVMVWVTDGESH-DGSMUKAVIDQCNHNDILRFGIAGLVGLNRLNALTGN 324
QY 280 SRQELNTIASKPPRDHVFQVANNFEALKTIONQLBEKIPAIEGTQTGSSSEHEHMSGQF 339
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 325 LIKEIKAIASIPTRYFFNFVSDAALLEKAGTLGEQIFSIETVQG--GDNFQEMSQYGF 383
QY 340 SAAITSNGP--LLSTVGSYDWAGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYAAAI 394
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 384 SADYSSQNDILMLGAVGAFGSGTIVOKTSHGHLIPPKQAFDQILODRNHSYLGYSVAA 443
QY 395 ILNRVQSLVLGAPRYQHIGLVAMPQNTGMWESNANV-----KGTQIGAYFGASLCSV 448
Db 444 ISTGESTHFVAGAPRANVTGQIVLYSVN-----ENGNITVIQAHRGDQIGSYFGSVLCSV 498
QY 449 DVDSNGSTDLVLIGAPHYYEOTR--GGQVSVCPILPRGORARWQCDAVLYGSGQGWGRFG 506
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 499 DVDKDTITDVLVAGAPMYMSDLKKEBGRVYLFTIKKGILGQH--FLEGPEGIENTRFG 555
QY 507 AALTVLGVGVNGDKLTDVAIGAPGEDNRGAVLVPHGTSGSGISPSHSORIGAS--KLSPR 564
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 556 SAIAALSINDNGFNDVIVGPLENONGAVIYINGHGT--IRTKYSKILGSDGAPRSH 614
QY 565 LQYFQSLSGGQDLTMDGLDLTVGAQGHVLLLSQPVLRVKAIEMFNPREVARNVFECN 624
Db 615 LQYFGRSLDYGDLNGDSITDVSIGAFGVVQLMSQSIADVAIEASFTEKI--TLVNKN 672
QY 625 DVVKGKEAGEVRVCLHVQKSTRDLREGQIQSVVYTDALD-----SGRPHSRAVFNETH 680
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 673 AQII-----LKLCF-----SAKFRPTKQNNQVAIVYNTILDADGFSRVRTSRGLFKENN 721
QY 681 NSTRQTOVLGLTQTC--ETLKLQLPNCIEDVPSPVLRLNFSLVGTPLPSAFGRNLPVLA 738
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 722 ERCLQNMVNVQAQSCPBHIIYIQPS---DVNSLDLURVDISLENPGTS-----PALE 772
QY 739 EDAQRLFTALFPFKNCGNDNICDDLSITF-----SFMSLDCLVYVGGPREFNVTVTRND 794
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 773 AYSETAKVSPFPHKDCGEDGLCISDLVDVRQIPAAQEQPIVSNQNKRLTFSVTLKXK 832
QY 795 GEDSYRTQVTFPFLDLDSYRKVSTLQNSORSORSLACESASST--EVSGALKSTSCSINH 853
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 833 RESAYNTGIVVDVDFSENLF-----ASFSLPVDGTEVTQVAAASQKSVACDVGY 880
QY 854 PIPPENSVEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTETFOLELPVKYAVMV 913
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 881 PALKREQVQVTFINPFDNLQ--NLQNASLSFOALSESQENKADNLVNLKIPLLDYDABI- 938
QY 914 VTSHGVSVKYNFNFTASENTSRVMQHOYQVSNLQOR-----SLPISLVFLV----- 958
Db 939 ---HLTRSTNINFEISSDGNVPSIVHSFEDVGPKFISLKVVTGCVSPVSMATVILHPQ 995
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 959 -----PVRLNQTVIWRPQVTF--SENLSSTCHTKER 988
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 996 YTKENKPLMYLTGVQTDKAGDISCNADINPLKIGT-----SSVSFKSENPR---HTKE- 1047
QY 989 LPSHSDFLAELRKAPVNCSTAVCORIQCDIPFFGIQDEFNATLKGNSLFDWYIKTSHNH 1048
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1048 -----LNCRTASCSNVTCWLKDVHMKGEYFNVVTVTRINWGTFFASTFTQT 1091

```

```

621 Db      PSGGDGKTLKPFQSIHGEMDLNGDGLTDVTIGLGGAAFWARDAVAVKVTMNFENKV 680
617 QY      ARNVPECNQVVKGEAG--EVRVCLHQV-KSTRDLREGQIQSVVTYDLALDSGRPHSR 678
681 Db      NIQKNCR--VEGKETVCINATMCFHVKLSKEDSIYEADLQ-----YRVTLDSLRQISR 733
674 QY      AVFNET-----KNSTRQTVGLGTQTCETKLQIOLPNCI-----EDPVSPIVLR 718
734 Db      SFGSGTQERKIQRNITVRESE-----CIRHSPYMLDKHDFQDSVRVTL 776
719 QY      NFSLVGTPLSAFGNLRPVLAEADAQRLFTALPFPKNCNDNI CODDLSITPFSMSLCLV 776
777 Db      DFNLT-DENG-----PVLDDALPNSVHEH1PFAKDCGNKERCISDLTLNVSTTTEKSLLI 833
779 QY      VGGPRE-FNVTVTVRNDCGDSYRTQVTFPFLDLSYRKVSTLQNGRSQRMSRLACESASS 833
831 Db      VKSQDKFNVSILTVKNKGD SAYNTRTVQHSPLNIFSGIEIQD-----SCSN-- 880
838 QY      TEVSGALKSTCSINHP1PFENSEVTFNITFDVDSKASLGKLL-LKANVTSENMPRTN 899
881 Db      -----QNITCRVGYPLRAGETVTFK11FQFNTSHLSENAIHLSATSDSESPLES LN 933
897 QY      KTEFOLELPLVKAV---YMVVTHSGVST-----KVLNFTASENTRSVMOHQYQVSNL 945
934 Db      DNEVNISIPVKYEVGLQFYSSASEHH1SVAANETIPEFINST--EDIGNEINVFTYTKR 999
946 QY      GQSRSLP--ISLVF-----LVPVRLNQTVWD-----RP-----Q 977
992 Db      GHFHPPELQLSISFNLTADGVPVLYPIG-----WSSSDNVNCRPSLEDPFGINSKK 1000
973 QY      VTFS-----ENLSTCHTKERLPSPSHDSFLAELKAPVWNCIAVCORIQC DI--- 1000
1046 Db      MTISKSEVLKRGTIQDCSSTC-----GVATITCSLLPSDLSQVNVSL 1000
1020 QY      ---PFGIOBEF---NATLKNLSFDWIKTSHNHLIVSTABILFNDSVFTLLPGQAF 1000
1089 Db      LMKPTF-IRAHFSSNLTLRGELK-----SENSSLTSSN----- 1100
1074 QY      VRSQETKVEPPEVNP1PL--IVGSSVGGLLLLALITAAALYKLGFFKRYQKDMWSE 1128
1124 Db      RRRELAIQISKDGLGRVPLWILLUSAFAGULLMLLILALWKIGFFKRP1KKWMEK 1180

RESULT 12
A41131
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N;Alternate names: integrin alpha-4
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 24-Sep-1999
C;Accession: A41131, S16742
R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmanna, B.; Weiseman, J.
J. Cell Biol. 115, 1149-1158, 1991
A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine I
A;Reference number: A41131; MUID:92064645; PMID:1840602
A;Accession: A41131
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1039 <NEU>
A;Cross-references: EMBL:X53176; NID:g51484; PIDN:CAA37316.1; PID:g51485
C;Superfamily: integrin alpha-4 chain
C;Keywords: cytoskeleton; transmembrane protein

Query Match      11.3%; Score 663; DB 2; Length 1039;
Best Local Similarity 22.7%; Pred. No. 2.3e-37;
Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 4

QY      1 FNLDTENAMTFQ-ENARFGGQSVW-QLOGSR--VVVGAPQEIIVAAN-----QRGSLYQCDY 52
Db      41 YNLDPENALLYQPGSTLFGYSVVLHSGSKRWLVIGAPTASWLSNASNVAWNPAGIYRCGI 100
QY      53 STG---SCPEIRLQVP-----VEAVNMSLGLSLAATTSP-POLLACG---PTVHQ 95

```

Db 101 RKNPNQTCLOLQSGSPGCGKTCLEERDQNLGVTLSRQPGENGSI VTCGHRWKNIFY 160
QY 96 TSENTRYVGLCLFGLFSGNLRQOPQKPFPEALRGCPQEDSDIAFLIDGSGSII PHDFRRMKE 155
Db 161 MKSDNKLPTGICVMPDLRTLSK----- 185
QY 156 FVSTVMEQLKSKTLFSLMOYSEEFRIHTFKFQNNPNRSLVPI TOLLGRTHTATGV 215
Db 186 -----RMAPCYKDYT----- 195
QY 216 RKVIRELLNITNGARKNAKILIVITDGEKFGDPLGYEDVIP EADREGVIRYVIGVDAF 275
Db 196 -----RKFGEE----- 200
QY 276 RSEKSRQELNTIASKPPRDHVFQVNNFEALKTIONLREKIFA IEGTQTGSSSSPEHEMS 335
Db 201 -----NFAS-----C 205
QY 336 QGFSAAITNSGNPLSTVGSYNWAGGVFLY--TSKEKSTFINM TRVDSMDNDAYLGYA-- 391
Db 206 QAGISSFYTDLIVMGAPGSSYMTGTVFVYNTTQYKAFVD--R QNVKFGSYLGYSVG 263
QY 392 AAILLRNVQSLVGLAPRYQHIGLVMFRONTGMWESNANV--K GTOIGAYFGASLCS 447
Db 264 AGHFRSPHTEVVGGAPQHEQIGKAYIF----SIDENELN IYVMKKGKLSYFGASVCA 319
QY 448 VDVDSNGTDLVLIGAPHYVEQTRGGQSVCPPLRGORA--R WQCDALVYGEQGPWGRFG 506
Db 320 VDLNADGFSDL-LVGAPMSTIREGRVFVY-INSNGAVVM EMERVLVGS DKYA-ARGF 376
QY 507 AALTVLGVNGDKLTDVAIGAPEEDNRGAVYLFHGTSGGSI SPSSHORIAGSKLSPRIQ 566
Db 377 ESIANLGDINDGFEIDAGQEDDLRGAVIYNGRV-DGISST YSQRIEQQISKSLR 435
QY 567 YGQSLSGQDLTMDGLVDLTGVA--QGHVLLRSQPLVRKAI MEFNPREVARNVFECN 624
Db 436 MFGQSISGQIDADNNGYGVDAVGAFQSDSAVLLRTPVVI VEASLS-HPESVNRTKFDC 494
QY 625 DQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVTDIALDS GR---PHSRAVF--NBT 679
Db 495 -----ENGLPSVCMHLTLCFSYKGEVPGYIVLFYNS LVDVHRKAESPFPFSNGT 547
QY 680 KNSTRQTOVLGTTQTCETLKLQLPNCIEDPVSPIVLRNFS L-----VGTPLSAFNGLR 734
Db 548 SDVITGSI RVSSGEXKCRTHQAFMRKDVRLDILTPIHVEAT YHLGHVITKRNTEBPPLQ 607
QY 735 PULAEADAQR-LFTALFPPEKNCNDNICODDLSITFS----- FMSLDCLVYGGPREFN 787
Db 608 PILQOKKEDVIRKMINFARFCAVEN-CSADLQVSAKVGLP KYENKTYLAVGSMKTI ML 666
QY 788 TVTVRNDGDSYRTOVTFEFLDLSYRKVSTLQNRQSRWR LACESASSTVSGALKST 847
Db 667 NVSLFNAGDVAETTLNVQLPTGLYFIKILDUDEK----- QINCE---VTSSGLV-K 716
QY 848 SCSSINHPFPENSEVTNFTFDVDSKASLGNKLLKANVTSEN -NMPRTNKTEFOLELP 905
Db 717 ACSLGIYVDRLSRIDISLLDVSSLSRAHEDLSISVHASC ENEGELDQVRNVRTLTIP 776
QY 906 VKYAVVMVTSRGV--STKYLNFPTASENTSRVWQHO----- YQVSNLQORSIP-LISLVL 957
Db 777 LRYEV--MLTVHGLVNPSTSFVYSGSSENEPETCMAEKL NLTHTVINTGISMAPNVSVKIM 834
QY 958 VPVRLNQTVIMWRP--QVTFSENLSTCHTKE-----RLP SHSDFLAELR 1000
Db 835 VP---NSFLPQDDKLFNVLDVQTTTCQCHFGYGRECTFA QQKGIAGTLTDIVKFLSKTD 891
QY 1001 KAPVWNCISIA--VCQRIQCIDFPFGIQEEFNATLKNLS FDFWYIKTSHNHLITVSTAEIL 1058
Db 892 KR-LLCYMKADQHCFLDN---FGKWESS-----KEASVHI QLEGPSL 933
QY 1059 FNDSVFTLLPGOGAFVRSQETKVEPFEVFPNP----- 1090
Db 934 EMDETSSL-----KFEIKATAPPEPHPKVIELNKDENV AHVFLGLEHHRKPRHF 983

QY 1091 -LPLIVSGSVGGLLLALITAAALYKLGFFKRYQKDMGSE 1128
Db 984 TIIITITISLLGLVILLISCVMMKAGFFKRYKSILQES 1022

RESULT 13

S06046
integrin alpha-4 chain precursor - human
N:Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C:Species: Homo sapiens (man)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 24-Sep-1999
C:Accession: S06046; A39355; D28018
R:Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other int
A:Reference number: S06046; MUID:89356603; PMID:2788572
A:Accession: S06046
A:Molecule type: mRNA
A:Residues: 1-1038 <TAK>
A:Cross-references: GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1; PID:g33946
R:Roehn, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A:Title: Characterization of the alpha-4 integrin gene promoter.
A:Reference number: A39355; MUID:91239513; PMID:2034655
A:Accession: A39355
A:Molecule type: DNA
A:Residues: 1-93 <ROS>
A:Cross-references: GB:M62841
R:Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A:Title: The very late antigen family of heterodimers is part of a superfamily of mole
A:Reference number: A94151; MUID:87204112; PMID:3033641
A:Accession: D28018
A:Molecule type: protein
A:Residues: 40-50, 'E', 52-53 <TA2>
C:Genetics:
A:Gene: GDB:ITCA4; CD49D
A:Cross-references: GDB:128032; OMIM:192975
A:Map position: 2q31-2q32
C:Superfamily: integrin alpha-4 chain
C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; trans
F:1-39/Domain: signal sequence #status predicted <SIG>
F:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 633; DB 2; Length 1038;
Best Local Similarity 25.2%; Pred. No. 2.8e-35;
Matches 248; Conservative 155; Mismatches 370; Indels 212; Gaps 39;

QY 272 GDAFRSEKSRQELNTIASKPPRD-----HVFQVNNFEALKT-----IQN 310
Db 121 GKTCLLEERDQNLGVTLSRQPGENGSI VTCGHRWKNIFYKKNKLP TGGCYGVPPDLRT 180
QY 311 QUREKI-----FAIEGTQTGSSSSPEHEMSQEGFSAAITSN GPLLTVGSY 356
Db 181 ELSKRIAPCYQYVKKFGENFA-----SCQAGISSFYTKDLI VMGAPGSS 225
QY 357 DWAGGVFLY--TSKEKSTFINMTRVDSMDNDAYLGYA--AAIL LRNVQSLVGLAPRYQ 411
Db 226 YMTGSLFVYNTITNKYKAPLDKQNKVF--GSLYGVSVGAGHFR SQHTTEVVGGAPQHE 282
QY 412 HIGLVMFRONTGMWESNANV---KGTQIGAYFGASLCSVDVDS NGSSTDVLIGAPHYV 467
Db 283 QIGKAYIP----SIDEKELNIIHEMKKGLSGYFGASVCAVD LNADGFSDL-LVGAPMOS 337
QY 468 EOTRGQVSVCPPLRGQRRMQC--DAVLYGEOQGPWGRFGAALT VLGDVNGDKLTDVAIG 526
Db 338 TIREGRVFVY-INSNGAVVMAMETNLVGS DKYA-ARFGESIV NLGDIINDGFEIDAG 395
QY 527 APGEEDNRGAVYLFHGTSGGSI SPSSHORIAGSKLSPRIQY GQSLSGQDLTMDGLVDL 586
Db 396 APQEDDLOGAIYIYNGRA-DGISSTFSORIEGLSKLSLMSF QSISGQIDADNNGYVDV 454

A:Description: Alpha SU2, a sea urchin integrin which binds laminin.

A:Reference number: Z21035

A:Accession: T31437

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1041 <HER>

A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1

A:Experimental source: developmental stage embryo

C:Function:

A:Description: binds laminin

C:Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 5.4e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY	336	QEGFSAAITSNCP--LLSTVGSYDWDAGGVFLYTSKEKSTFINNTRVDS-----DMNDAYLG 389
DB	181	QAGFGIIFSDNSALVNGAPGSYYLQGGIYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239
QY	390	YAAAI--ILNRVQSLVLGAPRYQHI-GLVAMERQNTGMWESNANVKGTOIGAYFGASLC 446
DB	240	YSLALGDFNGDVQDVTGTPRAESLMGLVAIFDQNLNQFN---QWGTQIVAYFGYSVT 296
QY	447	SVDDVDSNGSTDLVLIGAPHYEQTRGGQVSVCPRLPRQGRARWQCDVLYGEQ-----498
DB	297	VVDI--NNDVTYDDLVLGAPMTDGGPAIQ-----RWEAGAVVYVQLNPDPVGPQA 343
QY	499	-----QPNGRFGAALTVLGDVNGDKLTVAGAPCEENRGAVYLFHGTSGS 546
DB	344	SNRLSLSSTLIGQOIRSFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVIYHG--SAN 402
QY	547	GISPSHSQRIAGSKLS-PRLOYFGQSLSGQDITMTDGLVDTVGAQ--GHVLLLRQOPVL 603
DB	403	GLKSTPAQVLTPTLGHSGITTFGSLGGQDMKNKYPDLLVGAESANTAVLIRTPVV 462
QY	604	RVKAIMFNPREVARNVFECNDQVKGKEAGEVRVCLHVQKSTRDRLRREGIQSV-----658
DB	463	SLDALTLETP-----IGINLENKTYE-LADGTMVTSFIAMT 497
QY	659	-----VYYDLALDSG-RPHSRVFNETHNSTRROTQVIGL-TQTCETLKLQ 702
DB	498	CFTYTGNYLPDHDIDISVYTVVDSGIIANRRAMEFVNDMSIITKTRRLAVSTQFCDFLRAY 557
QY	703	LPNCIEDPVSPVLRNFSLVGTPLSAFGN-----LRPVLAEDAQRLLFTALPPEK 753
DB	558	VGNSIEDKLTPIKVTIQYDL-----NDESRLQPHETLPIIDMATMSTQTKQVSIQN 609
QY	754	NCGNNDICQDLSITFSFMSLDCLVGGPREFNVTVTRNDGEDSYRTQVTFPPFLDLSY 813
DB	610	NCVN-NICIPDLDTVT-PNLNPIVIGTQELTLDVSLNRRGEDAFQSSLSVYVPLGLQF 667
QY	814	RKVTIQNQRQSRWLACESASTVSGALKSTCSINHPIFPEN-----SEVTFNIT 867
DB	668	VRL-----ERKANMDFSVTCSDDSD-----LRIITCDTGNPMVGNILFGLTLSTFQVS 717
QY	868	FDVDS-----KASLGNKLLKANVTSNNMPRNKTEFQLELPVKYAVYVMTSHGVSTK 922
DB	718	GDKXDSIEFYFKAESSENS--EDPNTLENLNMVTPVTVDCITKLLSASYPEIVMYSTQED 775
QY	923	YL-----NFTASENTSRVMOHQYQV-----SNLQORSLPIS-----LVFLVPV 960
DB	776	YVVPFPKAKNASADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPKNKEDGEYLFYLLGI 834
QY	961	RLNQTVIWRPQ-----VTFSENLSSTCHTKERLPKSHSDFLAELRKAPVV 1005
DB	835	MTEEGVTCQLTQCKANPEGVKLEPSTKAKLSNSTTQVSGRKRREPEVAEALAQTDN--VI 892
QY	1006	NCSIAVCQRIQCDDIPFGIOEEFNAT-----LKNLSFDWYIKTSHNHLIIYSTAEILF 1059
DB	893	YCASDSCVLINCTI-----DEINASKSKVVRILGRF--W-----ERTF 928
QY	1060	NDSVFLLPGGAFVRSQTETKVE--PFEVPNP-----LP-----1092

Search completed: November 25, 2003, 14:22:09

Job time : 19.9579 secs

Db	929	QKAVSELTPVVOQTATIASASAAVKTIPYNIPLPRDFSDSTKASTLVTTEELVPPVPTPIAW 988
QY	1093	-LIVGSSVGGLLLLALITAAALYKLGFFKR-----QYKDMMS--EGGPP 1132
Db	989	WIIVSVVLGGIILLIILGLWKCGFFERKKPGEEKEYAPVASADKGGP 1038

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:09:04 ; Search time 9.17418 seconds
(without alignments)
5828.244 Million cell updates/sec

Title: US-09-902-481b-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFCQ.....FKRQYKDMMEGGPFCAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5852.5	99.6	1152	ITAM_HUMAN	P11215 homo sapien
2	4476	76.2	1153	ITAM_MOUSE	P05555 mus musculus
3	3473	59.1	1163	ITAX_HUMAN	P20702 homo sapien
4	3417	58.2	1162	ITAD_HUMAN	Q13149 homo sapien
5	1552.5	26.4	1170	ITAT_HUMAN	P20701 homo sapien
6	1533.5	26.1	1183	ITAL_MOUSE	P24063 mus musculus
7	1157.5	19.7	1167	ITAE_MOUSE	Q60677 mus musculus
8	1149	19.6	1179	ITAE_HUMAN	P38570 homo sapien
9	1102.5	18.8	1151	ITAI_HUMAN	P56199 homo sapien
10	1093.5	18.6	1189	ITAH_HUMAN	Q9UKX5 homo sapien
11	1084	18.5	1170	ITAZ_BOVIN	P53710 bos taurus
12	1072	18.2	1178	ITAZ_MOUSE	Q62469 mus musculus
13	1069	18.2	1181	ITAZ_HUMAN	P17301 homo sapien
14	1065.5	18.1	1167	ITAG_HUMAN	O75578 homo sapien
15	1060	18.0	1180	ITAI_RAT	P18614 rattus norv
16	663	11.3	1039	ITAA_MOUSE	Q00651 mus musculus
17	633	10.8	1038	ITAA_HUMAN	P13812 homo sapien
18	630	10.7	1035	ITAG_HUMAN	Q13797 homo sapien
19	593.5	10.1	1032	ITAA_XENLA	Q91687 xenopus lae
20	571.5	9.7	1066	ITAJ_CRISP	P17852 cricetidae
21	567.5	9.7	1053	ITAJ_MOUSE	Q62470 mus musculus
22	555.5	9.5	1053	ITAS_MOUSE	P11688 mus musculus
23	546.5	9.3	1050	ITAS_XENLA	Q06274 xenopus lae
24	543.5	9.3	1034	ITAV_CHICK	P23608 gallus gall
25	538.5	9.2	1130	ITAE_HUMAN	P23229 homo sapien
26	535	9.1	1044	ITAV_MOUSE	P43406 mus musculus
27	532	9.1	1049	ITAS_HUMAN	P08648 homo sapien
28	531.5	9.0	1066	ITAZ_HUMAN	P26006 homo sapien
29	530.5	9.0	1072	ITAV_CHICK	P26007 gallus gall
30	526	9.0	1048	ITAV_HUMAN	P06756 homo sapien
31	517	8.8	1044	ITAB_CHICK	P26009 gallus gall
32	512.5	8.7	1091	ITAE_MOUSE	Q61739 mus musculus
33	498	8.5	1179	ITAT_MOUSE	Q61738 mus musculus

RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098993; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

ALIGNMENTS

34	494	8.4	1396	1	ITA2_DROME	P12080 drosophila
35	491.5	8.4	1146	1	ITA1_DROME	Q24247 drosophila
36	489.5	8.3	1039	1	ITAB_HUMAN	P08514 homo sapien
37	489	8.3	1025	1	ITAB_HUMAN	P53708 homo sapien
38	489	8.3	1033	1	ITAB_MOUSE	O9QUM0 mus musculus
39	486	8.3	126	1	ITAM_CAVPO	P15578 cavia porce
40	472	8.0	1181	1	ITAT_HUMAN	Q13683 homo sapien
41	469.5	8.0	1106	1	ITAT_RAT	Q63258 rattus norv
42	466	7.9	1226	1	PAT2_CAEEL	P34446 caenorhabdi
43	445.5	7.6	1139	1	INAI_CAEEL	Q03600 caenorhabdi
44	423	7.2	1115	1	ITA3_DROME	Q44386 drosophila
45	383	6.5	1000	1	ITA5_DROME	Q9W1M8 drosophila

RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=920731318; PubMed=1683702;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
developmentally regulated expression";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=353202;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol:
conservation across species and homology to platelet IIB/IIIA";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
CR3 (CD11b/CD18)".
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
activation?".
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,
Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
assessment".
RL Structure 6:923-935(1998).
RN [12]
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
subunits and a calcium binding site on its lower surface".
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
RN [13]
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS
ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES
AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.
IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF
THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D
PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR
FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES
OF FIBRINOGEN GAMMA CHAIN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M
ASSOCIATES WITH BETA-2
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFWA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VFWA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD11b entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03925; AAA59544.1; -
DR EMBL; M18044; AAA59491.1; -
DR EMBL; J04145; AAA59903.1; -
DR EMBL; S22277; AAB24821.1; -
DR EMBL; S22152; AAB24821.1; JOINED.
DR EMBL; S22153; AAB24821.1; JOINED.
DR EMBL; S22154; AAB24821.1; JOINED.
DR EMBL; S22155; AAB24821.1; JOINED.
DR EMBL; S22157; AAB24821.1; JOINED.
DR EMBL; S22159; AAB24821.1; JOINED.
DR EMBL; S22161; AAB24821.1; JOINED.
DR EMBL; S22164; AAB24821.1; JOINED.
DR EMBL; S22165; AAB24821.1; JOINED.
DR EMBL; S22167; AAB24821.1; JOINED.
DR EMBL; S22169; AAB24821.1; JOINED.
DR EMBL; S22170; AAB24821.1; JOINED.
DR EMBL; S22173; AAB24821.1; JOINED.
DR EMBL; S22174; AAB24821.1; JOINED.
DR EMBL; S22180; AAB24821.1; JOINED.
DR EMBL; S22181; AAB24821.1; JOINED.
DR EMBL; S22184; AAB24821.1; JOINED.
DR EMBL; S22189; AAB24821.1; JOINED.
DR EMBL; S22191; AAB24821.1; JOINED.
DR EMBL; S22192; AAB24821.1; JOINED.
DR EMBL; S22203; AAB24821.1; JOINED.
DR EMBL; S22212; AAB24821.1; JOINED.
DR EMBL; S22213; AAB24821.1; JOINED.
DR EMBL; S22216; AAB24821.1; JOINED.
DR EMBL; S22219; AAB24821.1; JOINED.
DR EMBL; S22220; AAB24821.1; JOINED.
DR EMBL; S22221; AAB24821.1; JOINED.
DR EMBL; S22232; AAB24821.1; JOINED.
DR EMBL; S22236; AAB24821.1; JOINED.
DR EMBL; M76724; AAA58410.1; -
DR EMBL; M84477; AAA51960.1; -
DR PIR; A31108; RWHU1B.
DR PDB; 1A8X; 17-JUN-98.
DR PDB; 1BHO; 18-NOV-98.
DR PDB; 1BHQ; 18-NOV-98.
DR PDB; 1IDN; 25-NOV-98.
DR PDB; 1IDO; 01-AUG-96.
DR PDB; 1JLM; 11-JAN-97.
DR PDB; 1MIU; 07-AUG-02.
DR Genew; HGNC:6149; ITGAM.
DR MIM; 120980; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VFWA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; 3D-structure; Repeat; Magnesium; Calcium.
FT SIGNAL 1 16
FT CHAIN 17 1152 INTEGRIN ALPHA-M.
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1105 1128 POTENTIAL.
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).
FT REPEAT 31 84 FG-GAP 1.

Query Match		99.6%; Score 5852.5; DB 1; Length 1152;
Best Local Similarity		99.6%; Pred. No. 0;
Matches 1132; Conservative		3; Mismatches 1; Indels 1; Gaps 1;
QY	1	FNLDTENAMTFOENARGFGQSVVLOGSRVNVGAPQEI VAAANQSGSLYOCYSTGSCPEI 60
DB	17	FNLDTENAMTFOENARGFGQSVVLOGSRVNVGAPQEI VAAANQSGSLYOCYSTGSCPEI 76
QY	61	RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOK 120
DB	77	RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOK 136
QY	121	FPEARLGCQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQOLKSKTFLSLMQYSBEF 180
DB	137	FPEARLGCQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQOLKSKTFLSLMQYSBEF 196
QY	181	RTHFTPFKEFONNPNRSLVKPITOLLGRTHATGVRKVIRELLNITNGARKNAFKLIV 240
DB	197	RTHFTPFKEFONNPNRSLVKPITOLLGRTHATGVRKVIRELLNITNGARKNAFKLIV 256
QY	241	TGGEKFGDPLGYEDVPIEADREGVIRYVIGVGDAFSEKSRQELNTIASKPRDRHVQVN 300
DB	257	TGGEKFGDPLGYEDVPIEADREGVIRYVIGVGDAFSEKSRQELNTIASKPRDRHVQVN 316
QY	301	NFEALXTIQNLRKIFAJEGTQSGSSFEHEMSQEGFSAITSNGLLSTVGSYDWAG 360
DB	317	NFEALXTIQNLRKIFAJEGTQSGSSFEHEMSQEGFSAITSNGLLSTVGSYDWAG 376
QY	361	GVFLYTSKSKSFINMTRVDSMDNDAYLGAAAILRNVRQSLVILGAPRYQHIGLVAMFR 420
DB	377	GVFLYTSKSKSFINMTRVDSMDNDAYLGAAAILRNVRQSLVILGAPRYQHIGLVAMFR 436
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQSVVCP 480
DB	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQSVVCP 496
QY	481	PRGQARWQCDVLYGEOQPMGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB	497	PRG-RARWQCDVLYGEOQPMGRFGAALTVDGVDNGDKLTDVAIGAPGEEDNRGAVYLF 555
QY	541	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 600
DB	556	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGCAQGHVLLRSQ 615
QY	601	PVLVRKAIIMEFNPREVARNVFCNDQVVGKAGEVRVCLHVQKSTRDRRLRGQIQSVVT 660
DB	616	PVLVRKAIIMEFNPREVARNVFCNDQVVGKAGEVRVCLHVQKSTRDRRLRGQIQSVVT 675
QY	661	YDLALDSGRPHSRAVFNKSTRRTQVLGTLTQCTETLKLQLPNCIEDPVSPIVLRNF 720
DB	676	YDLALDSGRPHSRAVFNKSTRRTQVLGTLTQCTETLKLQLPNCIEDPVSPIVLRNF 735
QY	721	SLVGTPLSAFNGRPLVLAEDAQRFLTALPFPEKNCNDNICQDDLITFSFMSLDCLVVG 780
DB	736	SLVGTPLSAFNGRPLVLAEDAQRFLTALPFPEKNCNDNICQDDLITFSFMSLDCLVVG 795
QY	781	GPREFNVTVVRNDGSDSYRTQVTFPLDLSYRKVSTLQNRQSRQSWLACESASSTEV 840
DB	796	GPREFNVTVVRNDGSDSYRTQVTFPLDLSYRKVSTLQNRQSRQSWLACESASSTEV 855
QY	841	SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANYTSENMPRTNKTEF 900
DB	856	SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANYTSENMPRTNKTEF 915
QY	901	QLELPVKYAVYVMVTHSGVSTKYLNFTASENTSRVNHQVQVSNLQORSLSPLVLPV 960
DB	916	QLELPVKYAVYVMVTHSGVSTKYLNFTASENTSRVNHQVQVSNLQORSLSPLVLPV 975
QY	961	RLNQTIVMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1020
DB	976	RLNQTIVMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1035

1021 PFGIOEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQCGAFVRSQTB 1080

1036 PFGIOEBFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQCGAFVRSQTB 1095

1081 KYEPPEVNPPLIYGVSSVGGLLALLALITAAALYKLGFFKRYKDMWSEGGPPGASBP 1137

1096 KYEPPEVNPPLIYGVSSVGGLLALLALITAAALYKLGFFKRYKDMWSEGGPPGASBP 1152

RESULT 2

ITAM MOUSE

ID ITAM MOUSE STANDARD; PRT; 1153 AA.

AC P05555;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI).

GN ITGAM.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1] _TaxID=10090;

RP SEQUENCE FROM N.A.

RX MEDLINE=88312584; PubMed=3044779;

RA Pytela R.;

RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";

RT Willibrand factor.";

RL EMBO J. 7:1371-1378(1988).

RN [2]

RP SEQUENCE OF 11-45 FROM N.A.

RX MEDLINE=85188276; PubMed=3887182;

RA Springer T.A., Teplow D.B., Dreyer W.J.;

RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";

RL Nature 314:540-542(1985).

CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3b. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS. EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.

CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.

CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -!- SIMILARITY: Contains 7 FG-CAP repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X07640; CAA30479.1; -
DR EMBL; M14293; AAA39484.1; -
DR PIR; S00551; S00551.
DR HSSP; P11215; IABX.
DR MGD; MGI:96607; Itgam.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; VWF_A.
DR Pfam; PF00357; FG-GAP_3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Calcium; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 1153
FT DOMAIN 17 1105
FT TRANSMEM 1106 1129
FT REPEAT 1130 1153
FT REPEAT 31 84
FT REPEAT ? ?
FT DOMAIN 164 350
FT REPEAT 337 400
FT REPEAT 401 452
FT REPEAT 454 515
FT REPEAT 517 575
FT REPEAT 580 632
FT CA_BIND 465 473
FT CA_BIND 529 537
FT CA_BIND 592 600
FT SITE 1132 1136
FT DISULFID 66 73
FT DISULFID 105 123
FT DISULFID 654 711
FT DISULFID 770 776
FT DISULFID 999 1023
FT DISULFID 1028 1033
FT CARBOHYD 58 58
FT CARBOHYD 86 86
FT CARBOHYD 391 391
FT CARBOHYD 696 696
FT CARBOHYD 734 734
FT CARBOHYD 801 801
FT CARBOHYD 881 881
FT CARBOHYD 907 907
FT CARBOHYD 941 941
FT CARBOHYD 980 980
FT CARBOHYD 994 994
FT CARBOHYD 1022 1022
FT CARBOHYD 1045 1045
FT CARBOHYD 1051 1051
FT CARBOHYD 1076 1076
SQ SEQUENCE 1153 AA; 127480 MW; 178DB98AECB0343 CRC64;

Query Match 76.2%; Score 4476; DB 1; Length 1153;
Best Local Similarity 74.3%; Pred. No. 1e-294;
Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;
Qy 1 FNLDTENAMTFQENARFGQSVQLOGSRVVVGAPOEIVANORGLSYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFQENAKGFGQNVQLGGTGVVVAAPQAKAVNQTGALYQCDYSTSRCHPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHTQCSNTYVKGCLFLFGSNLRQPOK 120
Db 77 PLQVPEAVNMSLGLSLAVTVPQQLACGPTVHTQCKENTYVNGLCYLFSGNLLRPPO 136
Qy 121 FPEARLGCQEDSDIAFLIDGSGSIIPDPRFMKPFVSTMEOQLKSKTKLFLSLMQYSBF 180
Db 137 FPEARLGCQEDSDIVFLIDGSGSINNIDFQMKPFVSTMEOQLKSKTKLFLSLMQYSBF 196
Qy 181 RIHFTFKFQNNPNSLVPKIPITOLLGRTHATTGVKVIKRELLNTINGARKNAFKILVI 240
Db 197 RIHFTFNDPKRNPSPRSHVSPKIQKNGRTKTAGIKVVRKRELFHKTNGARENAKILVI 256
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELANTIASKPRDHVQVN 300
Db 257 TDGEKFGDPLDYKDVPEADRAGVIRYVIGVGNFKNPQSRRELDITASKPAGERHVQVD 316
Qy 301 NFEALKTIONQUREKIPAIETGTQSSSFEHMSQEGFSAITSNGPLSTVGSVDWAG 360
Db 317 NFEALNTIQLOEKIFAETGTQSTSSFEHMSQEGFSAITSNGPLSGVSGSPDWAG 376
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILRNVRQSLVLCAPYQHIGLVAMFR 420
Db 377 GAFLYTSKOKVFINTRVDSMDNDAYLGAAVILRNVRQSLVLCAPYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSDTLVLIGAPHYEEQTRGQGVSCPL 480
Db 437 ENFGTWEPTSIGSQIGSYFGASLCSDVMDADGNTNLILIGAPHYEEKTRGQGVSCPL 496
Qy 481 PRQARWQCDVLYGEOQOPWGRFGAALTVLGDVNGDKLTDAICAPGEEENRGAYVLF 540
Db 497 PRG-RARWQCEALLHGDQHPWGRFGAALTVLGDVNGDKLTDAICAPGEEENRGAYVIF 555
Qy 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSLGGODLTMDGLVDLTGCAQGHVLLRSQ 600
Db 556 YGASIASLSASHSHRIIGAHFSPGLQYFGOSLGGKDLTMDGLMDLAVAGQHLLLRAQ 615
Qy 601 PVLVRKAIMFEPNPREVARNVFECDQVVRKAGEVRVCLVHVKSTRDLRREGQIOSVVT 660
Db 616 PVLRLAETWEFSPKVARSVFACQEQVKNKADAGEVRVCLVRKNTKDLRREGDIQSTVT 675
Qy 661 YDLALDSGRPHSPAVNETNSTRROTQVLGLTQTCETLKLQPLNCIEDPVSIVLRNLF 720
Db 676 YDLALDPVRSIRAPFDETNNTRRTQVFLGQKCEKTLKILPDCVDSVSIILRLNY 735
Qy 721 SLVGTPLSAFGLNRPVLAEDAOFLTALPFEKNCGNDNICODDLSITFSFMSLDCLVWG 780
Db 736 TLVGEPLRSFGLNRPVLAEDAOQFFFTAMPFPEKNCGNDISICQDDLSITMSAGLDTLVWG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPLDLSYRKVSTLQNRQSQRWRL-ACESASSTE 839
Db 796 GPQDFNMSVTLRNDGEDSYGTQVTVVYPSGLSVRKDSASQNPITKXPMFKPAESSSSSE 855
Qy 840 VSGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSNNPRNTKTE 899
Db 856 GHGALKSTTWNINHPIFPANSEVFNITFDVDSHASFGNKLKLLKAI VASENNNSRTHKTK 915
Qy 900 FQLELPVKYAVVAVTSHGVSTKYLNFPTASENTRVMQHOYQVNLGORSPLISLVPLVP 959
Db 916 FQLELPVKYAIYVITSDSSIRYLNFTASEMSTKVIHQYOYFNNGLQVSLPVSVPWIP 975
Qy 960 VRLNQTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQICDI 1019
Db 976 VQINNVTVMDHPQVIFSNLSSACHTEQKSPHSHNFRDQLERTPVLNCSVAVCKRLOCDL 1035
Qy 1020 PFFGIGQEFNATLKNLSFDWYIKTSHNHLIIVSTABILFNDVSVFTLLPGQGFVRSQTE 1079
Db 1036 PSNTQEIFNVTLKGNLSPDWYIKTSHGHLVLSSTTEILFNDVSAFALLPGQSVRSKTE 1095
Qy 1080 TKVEPEVNPPLPLIVGSSVGGILLALLALITAAALYKLGFFKRYQKDMSEGGPGABEQ 1137
Db 1096 TKVEPEVNPPLPLIVGSSIGGLVLLALITAGLYKLGFFKRYQKDMSEAAPODAPQ 1153

```
RESULT 3
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
GN alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
DE ITGAX OR CD11c.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=8816645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of
a leukocyte adhesion glycoprotein, p150,95";
RL EMBO J. 6:4023-4028(1987).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte
p150,95 molecule";
RL J. Biol. Chem. 265:2782-2788(1990).
[3]
ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
[4]
SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
and p150,95 leukocyte adhesion proteins";
RL J. Immunol. 138:2381-2383(1987).
CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VFMA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
DR EMBL; M81695; AAA5180.1; -
DR EMBL; Y00093; CAA68283.1; -
DR EMBL; M29165; -; NOT ANNOTATED CDS.
DR EMBL; M29487; AAA51620.1; ALT_SEQ.
DR EMBL; M29482; AAA51620.1; JOINED.
DR EMBL; M29483; AAA51620.1; JOINED.
DR EMBL; M29484; AAA51620.1; JOINED.
DR EMBL; M29485; AAA51620.1; JOINED.

EMBL; M29486; AAA51620.1; JOINED.
DR PIR; A36584; RWHUIC.
DR PDB; 1N3Y; 18-FEB-03.
DR Genew; HGNC:6152; ITGAX.
DR MIM; 151510; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0004872; P:receptor activity; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1163 INTEGRIN ALPHA-X.
FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1108 1128 POTENTIAL.
FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).
FT REPEAT 34 87 FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT DOMAIN 165 351 VWFA.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 402 453 FG-GAP 4.
FT REPEAT 455 517 FG-GAP 5.
FT REPEAT 518 576 FG-GAP 6.
FT REPEAT 581 633 FG-GAP 7.
FT CA_BIND 466 474 POTENTIAL.
FT CA_BIND 530 538 POTENTIAL.
FT CA_BIND 593 601 POTENTIAL.
FT SITE 1131 1135 GFFKR MOTIF.
FT DISULFID 69 76 BY SIMILARITY.
FT DISULFID 108 126 BY SIMILARITY.
FT DISULFID 655 712 BY SIMILARITY.
FT DISULFID 771 777 BY SIMILARITY.
FT DISULFID 848 863 BY SIMILARITY.
FT DISULFID 998 1022 BY SIMILARITY.
FT DISULFID 1027 1032 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 490 490 G -> A (IN REF. 2).
FT CONFLICT 756 756 L -> D (IN REF. 2).
SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;

Query Match 59.1%; Score 3473; DB 1; Length 1163;
Best Local Similarity 61.3%; Pred. No. 7.8e-227;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGFGSGVVQLOGSRVVGAPQEIIVAAQGRGSLYQCDYSTGSCPEI 60
DB 20 FNLDTEELTAFRVDSAGFGSDVVGQYANSVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLIACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
DB 80 GLQVPPEAVNMSLGLSLASTTSPQLIACGPTVHCEGRNMYLTGLCFLLGPT--QLTOR 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTVMEQLKSKTFLSLMQYSEEF 180
DB 138 LPVSRQECPRQEQDIVFLIDGSGSISRRNFATMWNFVRAVISQFRPSTQFSLMQFSNKF 197
```


DB 1091 MVLEDEVYNAIPIMSSVGALLLALITATLYKLGFFRHYKEMLED 11139

RESULT 5

ITAL_HUMAN STANDARD; PRT; 1170 AA.

AC P20701; O43746;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Integrin alpha-L precursor (leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).

GN ITGAL OR CD11A.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.

RX MEDLINE=89139587; PubMed=2537322;

RA Larson R.S., Corbi A.L., Berman L., Springer T.;

RT "Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily.";

RL J. Cell Biol. 108:703-712(1989).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2)

RX MEDLINE=99425270; PubMed=10493829;

RA Loftus B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;

RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 1p and 16q.";

RL Genomics 60:295-308(1999).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.

RX MEDLINE=96036067; PubMed=7479767;

RA Qu A., Leahy D.J.;

RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin.";

RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.

RX MEDLINE=96398692; PubMed=8805579;

RA Qu A., Leahy D.J.;

RT "The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin.";

RL Structure 4:931-942(1996).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.

RX MEDLINE=99425288; PubMed=10493852;

RA Kallen J., Weizenbach K., Ramage P., Geyl D., Kriwacki R., Legge G., Cottens S., Weitz-Schmidt G., Hommel U.;

RT "Structural basis for LFA-1 inhibition upon lovastatin binding to the CD11a I-domain.";

RL J. Mol. Biol. 292:1-9(1999).

CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2, ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES AND MONOCYTES.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L ASSOCIATES WITH BETA-2.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=P20701-1; Sequence=Displayed;

CC Name=2;

CC IsoId=P20701-2; Sequence=VSP_002738;

CC Note=No experimental confirmation available;

CC

CC -1- TISSUE SPECIFICITY: LEUKOCYTES.

CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: Contains 1 VWFA domain.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry; WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)

CC

CC EMBL; Y00796; CAA68747.1; -.

DR EMBL; AC002310; AAC31672.1; -.

DR PIR; S03308; S03308.

DR PDB; 1LFA; 29-JAN-96.

DR PDB; 1ZON; 07-DEC-96.

DR PDB; 1ZOO; 07-DEC-96.

DR PDB; 1ZOP; 07-DEC-96.

DR PDB; 1COP; 07-AUG-00.

DR PDB; 1DQO; 03-FEB-00.

DR PDB; 1MQ8; 14-JAN-03.

DR PDB; 1MQ9; 14-JAN-03.

DR PDB; 1MQA; 14-JAN-03.

DR MIM; 153370; -.

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0006928; P:cell motility; TAS.

DR InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int_alpha; 4.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PS0234; VWFA; 1.

KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

KW Signal; 3D-structure; Magnesium; Calcium; Repeat;

KW Alternative splicing.

FT SIGNAL 1 25

FT CHAIN 26 1170 INTEGRIN ALPHA-L.

FT DOMAIN 26 1088 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1089 1112 POTENTIAL.

FT DOMAIN 1113 1170 CYTOPLASMIC (POTENTIAL).

FT REPEAT 42 91 FG-GAP 1.

FT REPEAT 92 149 FG-GAP 2.

FT DOMAIN 170 349 VWFA.

FT REPEAT 7 7 FG-GAP 3.

FT REPEAT 401 455 FG-GAP 4.

FT REPEAT 457 516 FG-GAP 5.

FT REPEAT 518 575 FG-GAP 6.

FT REPEAT 578 630 FG-GAP 7.

FT CA_BIND 468 476 POTENTIAL.

FT CA_BIND 530 538 POTENTIAL.

FT CA_BIND 590 598 POTENTIAL.

FT SITE 1115 1119 GFPR MOTIF.

FT DISULFID 73 80 BY SIMILARITY.

FT DISULFID 111 129 BY SIMILARITY.

FT DISULFID 653 707 BY SIMILARITY.

FT DISULFID 771 777 BY SIMILARITY.

FT DISULFID 845 861 BY SIMILARITY.

FT DISULFID 998 1013 BY SIMILARITY.

FT DISULFID 1021 1052 BY SIMILARITY.

FT CARBOHYD 65 N-LINKED (GLCNAC...) (POTENTIAL).

```

FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 730 730 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 862 862 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 885 885 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 897 897 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1060 1060 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC 954 954 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT Q -> QGVHGLVEMQTSKQILCRPAGDAEHTVGAQEGELPC
FT PWGSEAFRDNIRAGPCR (in isoform 2).
FT /FTid=VSP 002738.
FT R -> W (IN REF. 1 AND 2).
FT Y -> I (IN REF. 2).
FT 214 214 CONFLICT
FT 660 660 CONFLICT
FT 155 155 STRAND
FT 164 164 STRAND
FT 165 165 TURN
FT 169 169 HELIX
FT 186 188 TURN
FT 191 198 STRAND
FT 202 206 STRAND
FT 208 213 HELIX
FT 217 221 HELIX
FT 222 223 TURN
FT 229 229 STRAND
FT 233 243 HELIX
FT 244 244 TURN
FT 247 249 TURN
FT 250 250 TURN
FT 253 254 TURN
FT 256 263 STRAND
FT 274 276 HELIX
FT 277 278 TURN
FT 280 286 STRAND
FT 288 290 HELIX
FT 293 297 HELIX
FT 298 299 TURN
FT 300 302 TURN
FT 307 310 HELIX
FT 311 314 STRAND
FT 318 318 TURN
FT 319 328 HELIX
FT 329 330 TURN
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match
Best Local Similarity 34.5%; Pred. No. 7.5e-97; Length 1170;
Matches 405; Conservative 208; Mismatches 460; Indels 101; Gaps 37;

QY 1 FNLDTENAMTQ--ENARGFGQSVVQLQGRVVVGAFOEIVAAHQRSGLYQCDYSTGSCSE 58
Db 26 YNLDVRCARSFPPRAGRHFGRVLOV-GNGVIVGAPGE---GNSGSLYQCQSGTGCHL 81
QY 59 PIRLOVPEAVNMSLGLSLAATTSPPOLLACGPVHOTCSNTYVVGKLCFLGSLNLR--- 115
Db 82 PVTLR-GSNYTSKYLGLMTLATDPTDGSILACDPLGSLRTCDQNTYLSGLCYLFRNLQGM 140
QY 116 -QQPKPEALRGCPQSDITAFIDGSGSTIIPDFRRMKEFVSTVMEQLKSKTLFSLM 174
Db 141 LQRPFGQECIKG----NVLVFLFDGMSLQDFOKILDPMKDVWKLSNTSYQFAAV 196
QY 175 QYSEEFRIHFTKEFQNNPNSRLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAF 234
Db 197 QFSTSYKTEFDFSDYVKKRQDALLKHVGMLLTNTFGAINVYATEVFEELCARPDAT 256
QY 235 KILVITDGEKFGDPLGVEDVIPADREGVIRYIVGVDAFRSEKSRQELNTIASKPRPD 294
Db 257 KVLIIITDGE--ATDSGNIDAADK-----IIRYIIGIKGHFQTKESQETLHKFASKPASE 309
QY 295 HVFOVNNFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITSNGLLSTVG 354

```

```

Db 310 FVKILDTFEKLKDLTFTELQKKIYVIEGTSKQDLTSFNMELSSSGISADLSRGHAVNCAVG 369
QY 355 SYDWAGCVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAA-IILNRVQSLVLCAPRYQH 412
Db 370 AKDWAGGFLDLKADLQDDTFIGNELTPEVRAGLYGTVTWLPSSRQKTSLLASGAPRYQH 429
QY 413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQT 470
Db 430 MGRVLLFQBPQGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLIGAPLFVGEQ 489
QY 471 RGGQSVVCLPLPRQARARWQDAV--LYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAP 528
Db 490 RGRGRVIY-----QRRQLGFEVSELQDGPYPLGRFGEAITALTIDINGDLVDVAVGAP 544
QY 529 GREDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLMGLVDLTLV 588
Db 545 LEE--QGAVTFNRRHG-GLSPQSPQRIEQTQVLSGIQWFGRSIHGKDLLEGDLADVAV 601
QY 589 GAQGHVLLRSQPLRVKAIEMFNPREVARNPECNDQVY-KGKEAGEVRVCLHVKSTR 647
Db 602 GAESQMIVLSSRPVDMVTLMSPSPAEIPVHEVECSYSTSNKMKEGVNIITICFQI-KSLY 660
QY 648 DRLREGQIQSVTYDLALDSGRPHRAVFNKSTRQTVQLGLTQTCETLKLQLPNCI 707
Db 661 PQF-QGRLVANLTYTLQLDGHRTTRRGLFPGRHRLRNIAVT-TSMSCCTDFSFPFVCV 718
QY 708 EDPVSPVLRLNFSL---VCTPLS--AFGN-----LRPVLAEADAQRLFTALPPEKNCGN 757
Db 719 QDLISPINVSLNFSLWEEGTPRDQRAQKDIPIILRPSLHSETWEI-----PFEKNCGE 773
QY 758 DNICQDDLSITFSFMSLDCLVGCGPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVS 817
Db 774 DKKCEANLRVVSFSPARSRLRLTAFASLSVELSLNLEEDAYVWQLDLHFPGLSFRKVE 833
QY 818 TLQQRSSQSRWLACES--ASSTEVSQALKSTSCSINHPIFPENSEVTFNITFDVSKAS 875
Db 834 ML---KPHSQIPVSCBELPEESRLSRL---SCNVSSPIFKAGHSVALQMFNTLVNS 887
QY 876 LGNKLLKLLKANVTSENN---MPRTNKTEFOLELPVKYAVVMVTVSHCVSTKYLNTASEN 931
Db 888 WGSVELHANVTCNNEDSLLEDNSATTI---IPILYINILIQOQEDSTLYYSFTPKGP 944
QY 932 TSRVMOHQYOV---SNLQQRSLP-ISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTK- 986
Db 945 KIHQVKHMVQVRIQPSIHNDNIPTLEAVGVPOPPSEGPITHQWSVQMEPPV--PCHYED 1002
QY 987 -ERLPSHSD--FLAELRKAPVNVNCISIAVCORIOCDIPFGIQEFNATLKNLSFDWYIK 1043
Db 1003 LERLPDAAEPLCALFRCPVW-----FRQELVQVIGTLELVEIE 1044
QY 1044 TSHNHLIVSTABILNDSVFTLLPGCAFVRVQTKVPEFVFPNPLPLIVSSVCGLL 1103
Db 1045 AS-SMFSLCSLSISFNSSKHFLYGSNASL-AQVVMKVDVYVEKQMLYLVLSGIGGLL 1102
QY 1104 LLALITAAALKLGFKKQYKDMMSSEG-GPPGAEP 1136
Db 1103 LLLLIFVLVYKGVFFKRLNKEKMEAGRGVNGIP 1136

RESULT 6
ITAL MOUSE STANDARD; PKT; 1163 AA.
ID ITAL MOUSE
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (Leukocyte function associated molecule 1, alpha chain) (CD11a).
DN ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

NCBI_TaxID=10090;
[1] _
RN SEQUENCE FROM N.A. PubMed=2051027;
RP MEDLINE=91268576; PubMed=2051027;
RX Kaufmann Y., Iseng E., Springer T.A.;
RA "Cloning of the murine lymphocyte function-associated molecule-1
RT alpha-subunit and its expression in COS cells.";
RL J. Immunol. 147:369-374(1991).
[2]
RN SEQUENCE OF 24-42. PubMed=3887182;
RP MEDLINE=85188276; PubMed=3887182;
RX Springer T.A., Teplow D.B., Dreyer W.J.;
RA "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
RT glycoproteins and unexpected relation to leukocyte interferon.";
RL Nature 314:540-542(1985).
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
CC ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
CC AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L
CC SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED
CC LEUKOCYTES RECRUITMENT.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 1 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60778; AAA39426.1; -
DR PIR; I56126; I56126.
DR HSSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Magnesium; Calcium;
KW Repeat.
FT SIGNAL 1 23
FT CHAIN 24 1163
FT DOMAIN 24 1084
FT TRANSMEM 1085 1108
FT DOMAIN 1109 1163
FT REPEAT 39 88
FT REPEAT ? ?
FT DOMAIN 148 334
FT REPEAT ? ?
FT REPEAT 399 454
FT REPEAT 455 514
FT REPEAT 516 573
FT REPEAT 576 628
FT CA_BIND 466 474
FT CA_BIND 528 536

FT	CA_BIND	588	596	POTENTIAL.
FT	SITE	1111	1115	GFPR MOTIF.
FT	DISULFID	70	77	BY SIMILARITY.
FT	DISULFID	108	126	BY SIMILARITY.
FT	DISULFID	147	199	BY SIMILARITY.
FT	DISULFID	651	705	BY SIMILARITY.
FT	DISULFID	767	773	BY SIMILARITY.
FT	DISULFID	840	856	BY SIMILARITY.
FT	DISULFID	993	1009	BY SIMILARITY.
FT	DISULFID	1017	1048	BY SIMILARITY.
FT	CARBOHYD	86	86	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	185	185	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	668	668	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	724	724	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	776	776	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	880	880	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	890	890	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	899	899	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	927	927	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1056	1056	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	1163 AA;	128343 MW;	A7A3078489E8232F CRC64;

Query Match 26.1%; Score 1533.5; DB 1; Length 1163;
Best Local Similarity 34.0%; Pred. No. 1.4e-95;
Matches 398; Conservative 217; Mismatches 458; Indels 99; Gaps 36;

QY	1	FNLDPTNMTFOENA-RFGQSVVQLQSGRVVGAPOBIVAAQORGSLYQCDYSTGSC	59
Db	24	YNLDTRPTQSLAQAGRHFGYQVLIQEDG-VVVGAPGE---GDNTGGLYHCR	79
QY	60	IRLQVPEAVNMSGLSLAATSPOLLACGPTVHQCSTNTYKGLCFLGSNLRQ	119
Db	80	VSLH-GSNHTSKYLGMTLATDAKGLLACDPLGSLRTCDQNTYLSGLCYLP	138
QY	120	KPPALRCPCDSDIAFLIDSGSIIPHDPRMKFVSTVMEOLKSKTKTFLSMQ	179
Db	139	QNRPAYQCMKGKVDLVFLFGQSLSLRKDFEKMVDWKLNTSYQFAAVQF	198
QY	180	FRIHFTKEF-QNNPNRSLVKPITQLLGRTHATTGVKRVIRELLNITNGAK	238
Db	199	CRTEFTLDYVKQKNPDVLLGSVQPMFLLNTFRAINVVAHVFKESGARPDAT	258
QY	239	VITGKEGDFLGYEDVPIEADREG-----VIRYVIGVDAPFSEKSRQELN	291
Db	259	IIIDG-----EASDKGNISAAHDITRYIIIGKHGFVSQVKQTLHFASEP	304
QY	292	PRDHVFOVNFPEALKTIQNLREKIFAETGTQSGSSSFEHMSQEGFSAIT	351
Db	305	VEEFVKILOTPEKLDLTDLQRRYIALEGTRQDLTSFNWELSSGSLADLSK	364
QY	352	TVGSYDWAGGVF-LYTSKEKSTFTINMTRVSDMDNDAYLGIAAA-IILNR	409
Db	365	AVGAKDWAGGFLDLREDLQAGTFVQGEPLTSDVRGGYLGTYVAMWTSR	424
QY	410	YOHIGLVAMFR--QNTGMWESNANVKGTQIGAYFASLCSVDVDSNGSTDL	467
Db	425	YQHVQVLLFOAPEAGGRWNQTKIEGTQIGSYFGGELCSVDLDQGEAELL	484
QY	468	EQTRGGOVSVCLPRGORARWOCDAVLVGEQOPWGRFGAALTVLGDVNGDK	527
Db	485	GEQGRGVFTY---QRRQSLFEMVSELOQDPQYPLGRFGAALTALTDING	541
QY	528	PGEEDNRGAVLPHGTSGSGISPSHSQISAGSKLSPRLQYFQOQSLGGQDL	587
Db	542	PLEE--QGAIVIFNCKPG-GLSPQSPQRIQGAQVFPGRWFGSRTHGVKDL	598
QY	588	VGAQGHVLLRSQPVLRVKAIMEFNPVARNVFCNDQVVKGEAG-EVRVCLH	646

```

Db      599  VGEGRVVVLSRRPVVDVVTLSFSPBEIPVHEVECSYAREEQHGKVKACFRKPLT 658
Qy      647  RDRLRGQIOGVVYDLALDSGRPHSRAVFNETKSNTRRQVTLGLTQTCETLKLQLPNC 706
Db      659  PQ--FQGRLLANLSTYQLQDCHRMRSRGLPFDGSHLSGNTSITP-DKSCLDFFHFEPIC 715
Qy      707  IEDVPSIVLRNLSLV---GTPUSAFGN-LRPVLAEADAQLFTALPPFKNKNQDNICQ 762
Db      716  IQDLISPINVLSNLSLEEETPRDQGRAMQPIRPSIHIV-FKEIPFKNKGEDKKCE 774
Qy      763  DDLISITSPWLSLCLVVGPP-----REFNVTVVRNDGDSYRTQVTFPPFLDLSYRKV 816
Db      775  ANLTLSSPARS-----GURLMSSASLAVETWLSNSGEDAYWVRLOLDFPRGLSPKV 827
Qy      817  STLQQRQSRQSWRLACESASSTVSGAL-KSTSCSINHPIPPENSEVTFNITFDVDSKAS 875
Db      828  EMLO---PHSRMPVSCDEL-TEGSULLTKLKNVSSPIFKAGQEVSLQVMFNTLNS 882
Qy      876  LGNKLLKANVTSEN-NMPTNKTEFQLELPVKYAVVMVTVSHGVSKYLNFTASENTR 934
Db      883  WEDFVELNGTVHCENENSLOEDNSAATHIPVLYPVNLTKEQENSTLYISFTPKGPKTQ 942
Qy      935  VMQHQYQVNSLQGRSLPLSLVFLVPRLNQTVIDRPO----VTFSENLS-----TCHTK 986
Db      943  QVQHVYQV-----RIQPSAYDHNMT-LEALVGVPRPHSEDLITYTWSVQTDPLVTCHSE 996
Qy      987  E-RLPSSHDFLAELRKAPVNCSTAVCORIQDIPFFQIOEFENATLKNLSFDWYIKTS 1045
Db      997  DLKRPSE---AEQCLPGV-----QRCPIVF---RWEILIQVTGVELSKEIKAS 1042
Qy      1046  HNHLIVSTABILFNDVSFTLLPQOGAFVRSQTEKVEPFPVNPPLVIGVSSVGGLLLL 1105
Db      1043  -STLSLCSLSVSFNSSKHFLYGSKA-SEAQLVKVDLIEKEMLRVYVLSGIGGLVLL 1100
Qy      1106  ALITAALYKLGFFRQYKDMV-SEGGPGGAP 1136
Db      1101  FLIFLALYKVGFFKRNKLEKMEADGGVNGSP 1132

RESULT 7
ITAE_MOUSE
ID ITAE_MOUSE STANDARD; PRT: 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN ITGAE;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=AKR;
RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,
RA Kilshaw P.J., Wells J.H.;
RA "Murine M290 integrin expression modulated by mast cell activation.";
RL Immunity 1:393-403(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS

```

```

CC      WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC      -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC      -!- SIMILARITY: Contains 1 VMFA domain.
CC      -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC      or send an email to license@ebi.ac.uk).
CC      -----
CC      EMBL; U12236; AAC52142.1; -
CC      HSP; P11215; 1A8X.
CC      MGD; MGI:1298377; Itgae.
CC      InterPro; IPR000413; Integrin_alpha.
CC      InterPro; IPR002035; VMF_A.
CC      Pfam; PF01839; FG-GAP; 3.
CC      Pfam; PF00357; Integrin_A; 1.
CC      Pfam; PF00092; Vma; 1.
CC      PRINTS; PR01185; INTEGRINA.
CC      SMART; SM00191; Int_alpha; 3.
CC      SMART; SM00327; VMA; 1.
CC      PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC      PROSITE; PS50234; VMFA; 1.
CC      Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC      Signal; Repeat; Magnesium;
CC      Calcium.
CC      SIGNAL 1 19 BY SIMILARITY.
CC      CHAIN 20 1167 INTEGRIN ALPHA-E.
CC      CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.
CC      CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.
CC      DOMAIN 20 1164 EXTRACELLULAR (POTENTIAL).
CC      TRANSMEM 1115 1137 POTENTIAL.
CC      DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).
CC      REPEAT ? ? FG-GAP 1.
CC      REPEAT ? ? FG-GAP 2.
CC      DOMAIN 149 192 X-DOMAIN (EXTRA DOMAIN).
CC      DOMAIN 193 384 VMFA.
CC      REPEAT 449 501 FG-GAP 3.
CC      REPEAT 503 564 FG-GAP 4.
CC      REPEAT 566 631 FG-GAP 5.
CC      REPEAT 634 686 FG-GAP 6.
CC      CA_BIND 514 522 FG-GAP 7.
CC      CA_BIND 578 586 POTENTIAL.
CC      DOMAIN 185 191 POTENTIAL.
CC      SITE 1140 1144 GLU-RICH (ACIDIC).
CC      DISULFID 72 83 GFPR MOTIF.
CC      DISULFID 130 164 BY SIMILARITY.
CC      DISULFID 698 754 BY SIMILARITY.
CC      DISULFID 814 820 BY SIMILARITY.
CC      DISULFID 884 898 BY SIMILARITY.
CC      DISULFID 998 1023 BY SIMILARITY.
CC      DISULFID 1031 1047 BY SIMILARITY.
CC      CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 773 773 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 829 829 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 846 846 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 1013 1013 N-LINKED (GLCNAC. .) (POTENTIAL).

```

FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFD CRC64;

Query Match 19.7%; Score 1157.5; DB 1; Length 1167;
Best Local Similarity 29.0%; Pred. No. 4.1e-70;
Matches 358; Conservative 211; Mismatches 459; Indels 205; Gaps 43;

QY 1 ENLDTEA--MTFOENARGFGSGVVOLOGSRVVVVGAPQBIVAANORGS-----LYQCDY 52
DB 20 FNDVDMAVWVTLQPCAPAVLSLLHLDPSN-----NQCLLVARRSSNRNTAALYRCAL 74
QY 53 STGSCSEPIRLQVPEAVNMSLGLSLAAT--SPPOLLAC-GPTVHQTCSENTYVVKGLCFL 109
DB 75 SI-SPDEIACQ-PVEHICMPKRGYQGVTLVGNHGVLCVQVQAKFRSLNSELGTACSL 132
QY 110 FGSNLRQKQKPEALRG-----C-----PQS 131
DB 133 LTPNLDLQAAVPSDLGFLDPGCAHVDSGDYCRSKGSGTGEEKSARRRRVTEEBEED 192
QY 132 DSDIAFLIDSGSIIPHDFRMKEFVSTVMEQL--KSKTSLFSLMOYSEEPRIHFTPEF 189
DB 193 GTEIAIVLDGSGSIPGSDFOKKNFISTWNRNFEYKCFECFNALVQYGAIVQTEFDLOS 252
QY 190 QNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVITDGEKFGDP 249
DB 253 ROINASLAKVQSIQVQVKEVTKTASAMQVLDNIFIPSRGSRKALKVMVLDGIDFGDP 312
QY 250 LGYEDVIPADREGVIRYVIGVDAPFRSEKSEKQELNTIASKPPRDHVFQVNNFEALKTIQ 309
DB 313 LNLTTVINSPKMGQVVRFAIGVDRFKNNNTVRELKLIASDPKEATFKVTNYSALDGLL 372
QY 310 NQREKIPALEGTQTCSSSFHEMSEOGFSAAITSNP-LLSTVGSYDWAQGVLY-TS 367
DB 373 SKLQQRIVHMEGT--VGDALQYQLAQTFGSAQILDKGVLGTGAFNWSGALLYSTQ 429
QY 368 KEKSTFINMT-RVDS-DMNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFRONTQM 425
DB 430 NRGCCFLNQTKEDSRTVQSYLYGLSLAVLHKAHGISYVAGAPRHLKGAIVELKEDR- 488
QY 426 WESNA---NVKGTQIGAYFGASLCSDVDNSNGSTDVLVIGAPHYEQTRGGQVSCFLPR 482
DB 489 -EEDAFVRRIEGRQMGSYFGSVLCPVIDDMDGTDFLLVAAPFYHIRGEEGRVYVQVE 547
QY 483 GQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAP-----GEEDNRGA 536
DB 548 -QDASFSLAHTLSGHPGLTNSRFGFAVAAGVDINQDKFTDVAIGAPLEGFGAGDGASYS 606
QY 537 VYLFHGTSGSGISPSQRSIAGSKLSPRLQYFGSLSGGQDLTMDGLVLDLTVGAQGHVLL 596
DB 607 VYLYNGHSG-GLYDPSQDI RASVASGLHYFGMSVSGGLDFNGDGLADITVGSRSADV 665
QY 597 LRSQPLRVKAIEMFNPREVARNVFNCNDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
DB 666 LRSRPVVDLTVSMFTFP-----DALPMVFIGKM-DVNLCEFDSSVASEPGLREM 715
QY 654 QIGSVVTYDIALDSGRPHSRAVNEKYNSTRQTQVLGLTQTC-----696
DB 716 FLNFTVDVVDV-----TKQRQLQCEDSSGQCLRKWNKNGSFLCEHFVLI 760
QY 697 ETILKLQLPNCIEDPVSPILVLRNFSLVGTPLSAFGLNR-----PVLAEADAQLRPLALF--P 750
DB 761 STEEL-----CEEDCFSNITIKTYE-----FQISGGRDYPNTL--DHYKEPSAIFQLP 809
QY 751 FERNCGNDNTCODLSITFSFMSLDCLVVGPGPFNVTVTRNDEGDSYRTQVTFPPDL 810
DB 810 YEKDCKNKVFCIAEIQLTN-ISOQELVVGVTKEVTMNIISLTNSGDSYTMNMAVPRN 868
QY 811 LSVRKVSTLQNRQSRWRACSSASTEVSGALKTSKCSINHIPIIPENSEVTFNFTFDV 870
DB 869 LQFKKI-----QKPSFDPVQCDPKPV---ASVLVNMNKIGHPIIL-KRSSVNVSVTQWL 918
QY 871 DSKASUGNKLKLLKANVTSENNMPTNKTTEFQLELPVKYAVYVMVTVSHGVSTKVLNFTASE 930

DB 919 EESVFPNRTADITVTSNSNEKSLEARETR---SLQFRHAFIAYLSR--PSVMYMN--TSQ 971
QY 931 NTSRVMOHQVSNLQORSPLISLVFLVPVRLNQTIVDRPQVTFSENLSST-----CHT 985
DB 972 SPDSHKKEFFNFNGENLFGAVFQLOICVPIKLQDF-----QIVRVKNLTKTQDHTSCTQ 1025
QY 986 KERLPFSDHDLAEURLKAPVNVNCISIAVCQRIQCDDIPFFGIOBEFNATLKNLSFDWYIKTS 1045
DB 1026 SQEPACGSDPQVHKWHSVVCAI-----TSNKENVTVAAEISVG 1065
QY 1046 HNHLIIVSTA-----EILFNDSSVFTLLPGGCAVRSQTEKVEFF-----EVPNPLIV 1095
DB 1066 HTKQLLRDVSLEPLIGEISFNKSLYEGLNAB-----NHRKTITVIFLKEETSRLPLII 1119
QY 1096 GSSVGGLLLLALITAAALYKLGFFKQYQVDMSE 1128
DB 1120 GSSIGGLLVVVIILFKCGFFRKQYQNLLE 1152

RESULT 8
ITAE HUMAN
ID ITAE HUMAN STANDARD; PRT; 1179 AA.
AC P38570; Q9NZU9;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-1EL).
GN ITGAE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.
RC TISSUE=Leukemia, and Lymphocytes;
RX MEDLINE=94164962; PubMed=8119947;
RA Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B., Parker C.M.;
RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";
RL J. Biol. Chem. 269:6016-6025(1994).
RN [2]
RP REVISIONS TO 88-114.
RA Parker C.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 53-1179 FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=20138496; PubMed=10673275;
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G., Sholevaruk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;
RT "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";
RL Genome Res. 10:165-173(2000).
RN [4]
RP MUTAGENESIS OF ASP-109 AND PHE-316.
RX MEDLINE=20400502; PubMed=10837471;
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;
RT "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";
RL J. Biol. Chem. 275:25652-25664(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

QY 747 ALFPPEKNGNDNICODDLISITFSFMSLCLVGGPRENVTVTVVRNDEGDSYRTQVTF 806
Db 815 FOLPEKACKNKLFCVAELQLA-ITVSQOELVVGTLKELTLNINTSGEDSYMTSMALN 873
QY 807 FFLDLISYRKVSTLQNRORSRLACESASSTEVSGALKSTCSINHPFPENSEVTNFI 866
Db 874 YPRNLQ-----LKRQKPPSPNIQCDPQPV---ASVLIMCRIGHGHPVL-KRSSAHVSV 923
QY 867 TFDVSKASLGNKLLKANVTSENN-----MPTNKTEFO---LELPVKYAVVTVTSHQV 919
Db 924 VQOLEENAPNRTADITVTVTNSNERRSLANETHLQFRHGFVAVLSKPSIMVNTGOGL 983
QY 920 S--TKYLNFTASNTSRVQHOVQVSNLQORSPLISLFLVPLVRNQTIVMDRPOVTFSE 977
Db 984 SHKEFLFVHGEN---LFGAEYQ-----LQICVPTKLGLOVAAVKLRTQ 1028
QY 978 NLSSTCHTKERLPSSDFLAELRKAPVNVNCISAVCQRIQCDIPFGIOEFPNATLKGMLS 1037
Db 1029 ASVTCTWSQERACAYSS-VQHVVEHWSVSCVIA-----SDKENVTVAEIS 1073
QY 1038 FMYIKTSNHLIVST-----AEILFNDSVFTLLPGOGAFVRSQOTETKVPFVFPNPL 1091
Db 1074 WD-----HSELLKDVTELQILGEISFNKSLYEGNAENH--RTKITVFLKDEKYHSL 1125
QY 1092 PLIVSGVGLLLALITAAALYKLGFFKQYKDMHSE 1128
Db 1126 PIIKSGVGLLVILVILFKCGPFRKRYOOLNLE 1162

RESULT 9

ITAL_HUMAN STANDARD; PRT; 1151 AA.
ID ITA1_HUMAN
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit."
RL J. Biol. Chem. 268:2989-2996 (1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".
DR PIR; A45226; A45226.
DR PDB; 1QC5; 17-MAY-00.
DR Genew; HGNC:6134; ITGAI.
DR MIM; 192968; .
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; NAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA_A.

Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Repeat; Calcium; Magnesium; 3D-structure.
KW DOMAIN 1 1113
EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1137 1151
CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75
FG-GAP 1.
FT REPEAT ?
FG-GAP 2.
FT DOMAIN 147 360
VWFA.
FT REPEAT 349 404
FG-GAP 3.
FT REPEAT 405 457
FG-GAP 4.
FT REPEAT 459 520
FG-GAP 5.
FT REPEAT 540 599
FG-GAP 6.
FT REPEAT 602 654
FG-GAP 7.
FT CA_BIND 470 478
POTENTIAL.
FT CA_BIND 552 560
POTENTIAL.
FT CA_BIND 614 622
POTENTIAL.
FT SITE 1139 1142
GFFKR MOTIF.
FT DISULFID 54 64
BY SIMILARITY.
FT DISULFID 680 669
BY SIMILARITY.
FT DISULFID 675 728
BY SIMILARITY.
FT DISULFID 780 786
BY SIMILARITY.
FT DISULFID 850 858
BY SIMILARITY.
FT DISULFID 1002 1034
BY SIMILARITY.
FT DISULFID 1037 1044
BY SIMILARITY.
FT CARBOHYD 46 46
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074
N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085
N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F31AABF52808 CRC64;

Query Match 18.8%; Score 1102.5; DB 1; Length 1151;
Best Local Similarity 27.8%; Pred. No. 2.1e-66;
Matches 344; Conservative 211; Mismatches 487; Indels 195; Gaps 44;

QY 1 FNLDTENAMTFOENARG-FGQSVVQL---QGSRVVVVGAPQEIIVAAQNRGLSQCDYSTGS 56
Db 1 FNVVDKNSMTFSQPVEDMFGYTVQVYENEEGKVLIGSLVGOPKRTGDVYKCPVGRGE 60
QY 57 CSP-IRLQVQVBA-----VNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVYKGL 106
Db 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGGFLACGPLYAYRCGLHLYTTGI 119

DR	GO: 0004895; F:cell adhesion receptor activity; TAS.	
DR	GO: 0005518; F:collagen binding activity; TAS.	
DR	GO: 0007160; P:collagen-matrix adhesion; TAS.	
DR	GO: 0007517; P:muscle development; TAS.	
DR	InterPro: IPR000413; Integrin_alpha.	
DR	InterPro: IPR002035; VWF_A.	
DR	Pfam: PF01839; FG-GAP_3.	
DR	Pfam: PF00092; vwa; 1.	
DR	PRINTS; PR01185; INTEGRINA.	
DR	PRINTS; PR00453; VWFADOMAIN.	
DR	SMART; SM00191; Int_alpha_5.	
DR	SMART; SM00327; vwa; 1.	
DR	PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.	
DR	PROSITE; PS02334; VWFA; 1.	
KW	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;	
KW	Signal; Repeat; Calcium; Magnesium; Polymorphism.	
FT	SIGNAL 1 22	POTENTIAL.
FT	CHAIN 23 1189	INTEGRIN ALPHA-11.
FT	DOMAIN 23 1142	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1143 1165	POTENTIAL
FT	DOMAIN 1166 1189	CYTOPLASMIC (POTENTIAL).
FT	REPEAT 38 94	FG-GAP 1.
FT	REPEAT 102 163	FG-GAP 2.
FT	DOMAIN 167 345	VWFA.
FT	REPEAT 359 420	FG-GAP 3.
FT	REPEAT 422 475	FG-GAP 4.
FT	REPEAT 477 537	FG-GAP 5.
FT	REPEAT 539 598	FG-GAP 6.
FT	REPEAT 601 653	FG-GAP 7.
FT	DOMAIN 1154 1162	POLY-LEU.
FT	DOMAIN 1174 1177	POLY-ARG.
FT	CA_BIND 488 496	POTENTIAL.
FT	CA_BIND 551 559	POTENTIAL.
FT	CA_BIND 613 621	POTENTIAL.
FT	DISULFID 76 83	BY SIMILARITY.
FT	DISULFID 121 139	POTENTIAL.
FT	DISULFID 129 159	POTENTIAL.
FT	DISULFID 659 668	BY SIMILARITY.
FT	DISULFID 674 729	BY SIMILARITY.
FT	DISULFID 781 787	BY SIMILARITY.
FT	DISULFID 881 893	BY SIMILARITY.
FT	CARBOHYD 82 82	N-LINKED (GLCNAC. .)
FT	CARBOHYD 95 95	N-LINKED (GLCNAC. .)
FT	CARBOHYD 291 291	N-LINKED (GLCNAC. .)
FT	CARBOHYD 331 331	N-LINKED (GLCNAC. .)
FT	CARBOHYD 358 358	N-LINKED (GLCNAC. .)
FT	CARBOHYD 449 449	N-LINKED (GLCNAC. .)
FT	CARBOHYD 462 462	N-LINKED (GLCNAC. .)
FT	CARBOHYD 528 528	N-LINKED (GLCNAC. .)
FT	CARBOHYD 642 642	N-LINKED (GLCNAC. .)
FT	CARBOHYD 694 694	N-LINKED (GLCNAC. .)
FT	CARBOHYD 857 857	N-LINKED (GLCNAC. .)
FT	CARBOHYD 894 894	N-LINKED (GLCNAC. .)
FT	CARBOHYD 973 973	N-LINKED (GLCNAC. .)
FT	CARBOHYD 1032 1032	N-LINKED (GLCNAC. .)
FT	CARBOHYD 1040 1040	N-LINKED (GLCNAC. .)
FT	VARIANT 433 433	V -> M.
FT	VARIANT 524 524	/FTId=VAR_009889.
FT	VARIANT 972 972	R -> L.
FT	VARIANT 1003 1003	/FTId=VAR_009890.
FT	VARIANT 1030 1030	L -> P.
FT	VARIANT 1030 1030	/FTId=VAR_009891.
FT	VARIANT 1030 1030	I -> M.
FT	VARIANT 1094 1094	/FTId=VAR_009892.
FT	VARIANT 1094 1094	MISSING.
FT	VARIANT 1094 1094	/FTId=VAR_009893.
FT	VARIANT 1094 1094	L -> V.
FT	VARIANT 1094 1094	/FTId=VAR_009889.
FT	VARIANT 1094 1094	/FTId=VAR_009894.
FT	VARIANT 1094 1094	/FTId=VAR_009895.
FT	VARIANT 1094 1094	/FTId=VAR_009896.
FT	VARIANT 1094 1094	/FTId=VAR_009897.
FT	VARIANT 1094 1094	/FTId=VAR_009898.
FT	VARIANT 1094 1094	/FTId=VAR_009899.
FT	VARIANT 1094 1094	/FTId=VAR_009900.
FT	VARIANT 1094 1094	/FTId=VAR_009901.
FT	VARIANT 1094 1094	/FTId=VAR_009902.
FT	VARIANT 1094 1094	/FTId=VAR_009903.
FT	VARIANT 1094 1094	/FTId=VAR_009904.
FT	VARIANT 1094 1094	/FTId=VAR_009905.
FT	VARIANT 1094 1094	/FTId=VAR_009906.
FT	VARIANT 1094 1094	/FTId=VAR_009907.
FT	VARIANT 1094 1094	/FTId=VAR_009908.
FT	VARIANT 1094 1094	/FTId=VAR_009909.
FT	VARIANT 1094 1094	/FTId=VAR_009910.
FT	VARIANT 1094 1094	/FTId=VAR_009911.
FT	VARIANT 1094 1094	/FTId=VAR_009912.
FT	VARIANT 1094 1094	/FTId=VAR_009913.
FT	VARIANT 1094 1094	/FTId=VAR_009914.
FT	VARIANT 1094 1094	/FTId=VAR_009915.
FT	VARIANT 1094 1094	/FTId=VAR_009916.
FT	VARIANT 1094 1094	/FTId=VAR_009917.
FT	VARIANT 1094 1094	/FTId=VAR_009918.
FT	VARIANT 1094 1094	/FTId=VAR_009919.
FT	VARIANT 1094 1094	/FTId=VAR_009920.
FT	VARIANT 1094 1094	/FTId=VAR_009921.
FT	VARIANT 1094 1094	/FTId=VAR_009922.
FT	VARIANT 1094 1094	/FTId=VAR_009923.
FT	VARIANT 1094 1094	/FTId=VAR_009924.
FT	VARIANT 1094 1094	/FTId=VAR_009925.
FT	VARIANT 1094 1094	/FTId=VAR_009926.
FT	VARIANT 1094 1094	/FTId=VAR_009927.
FT	VARIANT 1094 1094	/FTId=VAR_009928.
FT	VARIANT 1094 1094	/FTId=VAR_009929.
FT	VARIANT 1094 1094	/FTId=VAR_009930.
FT	VARIANT 1094 1094	/FTId=VAR_009931.

47:


```

Db 370 VGFSAEYSPQNNILMGAVGAYDWSGTVVQKTPHGHILFISKQAFQILQDRNHSYLGYS 429
Qy 392 AAILLRNVQSVLVCAPRYQHILGLVAMFRONTGMWESNANV-----KGTOIGAVFGASL 445
Db 430 VASISTGNSVHFVAGAPRANYTGQILVSVN-----ENGNTVTVIOSQGDQIGSYFGSVL 484
Qy 446 CSVDVDSNGSTDLVLIGAPHYEQTR--GGQVSVCPPLRGORARQWCDAVLYGEGQCPWG 503
Db 495 CAVONKOTITDVLVGLAPMYNDLKKEGRVYLFITKG-ILNWH--QFLEGNPLENA 541
Qy 504 RFGAALTVLGVNGDKLTDVAIGAPGEDNRGAVLYFHTSGSGISPSHSORIAS--KL 561
Db 542 RFGSAIALSDINMGDFNDVIVGSPLENQNSGAVIYNGHEGM-IRLRSYKILGSDRAF 600
Qy 562 SPRLOYFGOSLGGODLTWDGLDVLTVGAQGHVLLLRQPVLRVKAIMEFNPVARNVF 621
Db 601 SSHLOYFGSLDGYDGLNGSDITDVSVAFGVQVQLWQSOSIADSVDSFPPKIL--TIL 658
Qy 622 ECNDQVVGKEAGEVRVCLHVQKSTRDLRREGQIQSVVTVYDIALD-----SGRPHSRAVFN 677
Db 659 NKNAEI-----KLKCF-----SAKFRTNQNNQVAIVYNTIDEDQFSSRVISRGFLK 707
Qy 678 ETKNSTRQTQVLGLTQTC--TLKLQPLNCIEDPVSPIVLRNPSL--VGTPLSAFNL 733
Db 708 ENNERCLOKTMIVSQAQRCEVILHIEPS---DIISPLNLCWNISLENPGT-----756
Qy 734 RVLAEADQRLTALFPFENKCNNDNICODLSITF-----SFMSLDCLVVGPRFNVTV 789
Db 757 NPALAYSATKVSFIPFHKDCGDDGVCSIDLVLNVQQLPATQQPFIVSNQKRLTFSV 816
Qy 790 TVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASST-EVSGALKSTS 848
Db 817 QLNKKEAYNTEIUVDFSENLF-----ASHMPVDGTEVTCIASSQKSVT 864
Qy 849 CSINHPFPENSEVETNITFDVDSKASLKNKLLKANVTSENMPRTNTEFOLELPVKY 908
Db 865 CNVGVPAKSKQGVTFITNFENLQ-NLQNASISPRALSQENMADNSVNLKSLLY 923
Qy 909 ANYVWVTHGVSTKYNLFTASNTSRVMOHQVSNLGR-----SLPISLVFLV 958
Db 924 DAEIHIT-RSTNINFEVSLDGNVSVV-HSFE-DIGPKFTFSIKVTGVSVPVNSA---976
Qy 959 PVRLNQTIVDRPQVTFSEN---LSSTCHTKE-----RLPSHSDFLAR- 998
Db 977 -----SVLIHPQYTKRNLPLMYLTVHTDQAGDISCAEINPLKIGTSSSVFPSKEN 1030
Qy 999 LRKAPVNVGSIACQRIQCDIPFGIOEBFNATLKNLSFDWYIKTSHNHLIVSTAEI- 1057
Db 1031 FRHIKELNCRKTASCSNIMCWLRLQVKEVFLNVSTRWNGTFAASTFOTVQLTAAAEID 1090
Qy 1058 LFNDSVFTL-----LPGQAFVRSQTEKVEPFE-VPNPLPLIVGSSVGGLLLLALITA 1110
Db 1091 TYNPQIVVTEENTVITP-----LTIKPKHEKVEVPTGVIGSVIAGILLLLALVA 1140
Qy 1111 ALYKLGFFKRYOKDM 1125
Db 1141 ILKLGFFFRKRYOKM 1155

RESULT 12
IT2A MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Danjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VFMA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 229987; CAA82877.1; -
CC EMBL; X75427; CAA53178.1; -
CC PIR; S44142; S44142.
CC HSP; P17301; IAOX.
CC MGD; MGI:96600; Itga2.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 4.
CC SMART; SM00327; VWA; 1.
CC DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC DR PROSITE; PS0234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Platelet; Signal; Repeat; Calcium; Magnesium.
CC SIGNAL 1 26
CC CHAIN 27 1178
CC DOMAIN 27 1129
CC TRANSMEM 1130 1151
CC DOMAIN 1152 1178
CC REPEAT 42 100
CC REPEAT 185 375
CC DOMAIN 431 483
CC REPEAT 485 546
CC REPEAT 548 607
CC REPEAT 612 664
CC CA_BIND 496 504

```

FT CA_BIND 560 568 POTENTIAL.
 FT CA_BIND 624 632 POTENTIAL.
 FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1154 1158 GPRK MOTIF.
 FT DISULFID 80 89 BY SIMILARITY.
 FT DISULFID 677 734 BY SIMILARITY.
 FT DISULFID 786 792 BY SIMILARITY.
 FT DISULFID 862 873 BY SIMILARITY.
 FT DISULFID 1016 1047 BY SIMILARITY.
 FT DISULFID 1052 1057 BY SIMILARITY.
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 696 696 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;
 Query Match 18.2%; Score 1072; DB 1; Length 1178;
 Best Local Similarity 28.1%; Pred. No. 2.5e-64;
 Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;
 Qy 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRYVVGAPQEIIVAANQSGSLYQC--DYST 54
 Db 27 YNVLGPKAIFSGPSSQFGYVQQLTNPQGNMLLVGSPWSPFPENRMGDVYKCPVDLPT 86
 Qy 55 GSECPRLQ-----VPVAVNMSLGLSLAATSPPELLACGTVHTQTSENTYVYKGLC 107
 Db 87 ATCEKLNLSASISNVTETKTNMSLGLTLTRNPGTCGFLTCGLPLWAHQCGNQYATGIC 146
 Qy 108 FLFGSLNRQQPO---KEPEALRGCPQEDSDIAFLIDGSGSIIDHDFRMKEFVSTWMEQL 164
 Db 147 ----SDVSPDFQFLTSPPAVQCPSL--VDVVVCDSDNSIYP--WEAVKNEFLVKFTGL 199
 Qy 165 K--KSKTLFSLMOYSEFRHFTFKFQNNPNRSLVKPITQLIG-RTHATGVRKVKIRE 221
 Db 200 DIGPKKTOVALIOYANEPRIIFNLNDFETKEDVMVQATSETRQHGDLTNTFRAIEFARDY 259
 Qy 222 LNLITNGARKNAKILVIDGEKFGDPLGYEDVIPADREGVIRYVIGV-----GDAFR 276
 Db 260 AYSQTSGRPGATKVMVWVTDGESH-DGSKLKTIVQCNDBEILRFGLVGLNRLNLD 318
 Qy 277 SEKSRQELNTIASPPRDHVFQVNNFEALKTIONLREKIFAEGTQSGSSFEHMSQ 336
 Db 319 TKNLKEIKAIATPTTERTYFFNVADEAALLEKAGTLGEQIFSGIEGTVOG-GDNPQMEMAQ 377
 Qy 337 EGFSA--AITSNGPLLSVGSYDAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
 Db 378 VGFSAADYAPQNDILMLGAVGAFDWSGLTVQETSHKPVIFPKQAFDQVLQDRNHSFLGYS 437
 Qy 392 AAILNRVOSLVGAPRYQHIGLVAMFRONTGMESNANV---KQTQIGAYFGASLCS 447
 Db 438 VAAIATEDGVHVFAGAPRANVTGQIVLYSVNK---QGNVTVIQSHRGDQGSYFGSVLCS 494
 Qy 448 VDSDNSGTDVLVIGAPHYEQTR--GGQVSVCPPLPGQARWQCDVAVLYGEOGPWGRF 505
 Db 495 VDVOKTITIVLLVAGTYNDLKEEGKYLFTITKGLNQHQ---FLEGPEGTGNARF 551
 Qy 506 GAALTVLGVNGDKLTDVAICAPEEDNRAVYLFHGTSGSGISPSHSQRIAGSKLSPR- 564
 Db 552 GSAIAALSDINMGDFNDVIVGSPVENENSGAVIYNCHQGT-ITKYSQKILSGNGAFRR 610
 Qy 565 -LQYFGSLGSGQDLTWDGLVDLTVGAGHVLLLRSQPVLVRVKAIMEFNPVARNVFEF 623
 Db 611 HLQFFGRSLGSLGDLNGDSITDVSIGALGVQVQLWSQSIADVAIEALFTE----- 660
 Qy 624 NDQVVKKEAGEVRVCLHVQKSTRDLREQIQSVVYTDALD-----SGRPHSAVENET 679
 Db 661 -DKITLLNKDAKITLKLCFRAEPRPQGNQV--AILFNMTLDADGHSRVRTSRGVREN 717

Qy 680 KNSTRQTVLGLTQTCET--LKQLPNCIEDPVSPVLRLNLSFSLVGTPLSAFQNLRPVL 737
 Db 718 SERFLOKMMVNEVQKCEHHSIQKPS---DVVNPLDLRVDISLENGPS-----PAL 768
 Qy 738 AEDAQRLFTALFFPKCNCGNDNICQDDLSI-----TFSPMSLDCLVVGPREPNVT 788
 Db 769 EAYSETVKVFESIPFYKECGSDGICISDLILDVQOLPAIQTSF-----IVSNQNKRLTFS 823
 Qy 789 VTVNRNDESYRTQVTFPFFPLDLSYRKVSTLQNRQSRWRLACESASST-EVSGALKST 847
 Db 824 VILKNGESAYNTVLAEFSENLF-----ASFMPVDGTEVTCVSGSQSV 871
 Qy 848 SCINHPFIPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTE--FQLELP 905
 Db 872 TCDGVGPALKSEQQVTFITNFDNLQ-NLQNAQAINFAQPSSEQ--ETNKADNSVSLTIP 928
 Qy 906 VKYAVYVVTSHGVSTKYLNTASENTSRVMQHQYQVSNLQOR-----SLPLSLV 955
 Db 929 LLYDABELHLT-RSTNINFIYSIDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVMA 984
 Qy 956 FLV-----PVRNLTQVWDRPQVTF--SENLS 980
 Db 985 LVTIHIPOYTKENKPNLLYLTGIQTDQAGDISCTAEINPLKPHTA----PSVFKQENFR 1040
 Qy 981 STCHTKERLPSSHDFLAELRKAPVNCIAVQCORIQCDIPFGIOBEFNATLKNLSFDW 1040
 Db 1041 ---HTKE-----LDCRTTSCNITCWLKDLHMAEYFINVTRVWNR 1080
 Qy 1041 YIKTSHNLLIVSTAILFNDSVFTLLPGOGAFVRSQTETKVEPFPVNPPLVIGSSVG 1100
 Db 1081 FAASTFTQVLTAAASIDITHNPOLFVIEENAVTIPLIMKPKTEKAEVPT--GVIIGSIIA 1138
 Qy 1101 GLLLALITAAVLKLGFFKQYKDM 1125
 Db 1139 GILLALLAMTAGLWKLGFRRQYKDM 1163
 RESULT 13
 ID ITA2 HUMAN STANDARD; PRT; 1181 AA.
 AC P17301;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
 GN ITGA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
 RC TISSUE=Endothelial cells;
 RX MEDLINE=89308879; PubMed=2545729;
 RA Takada Y., Hemler M.E.;
 RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit
 (platelet GPIa): homology to other integrins and the presence of a
 possible collagen-binding domain.";
 RT J. Cell Biol. 109:397-407 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Rieder M.J., Amel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
 RA Rajkumar N., Toth E.J., Yi O., Nickerson D.A.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
 RX MEDLINE=98019223; PubMed=9353112;
 RA Emelley J., King S.L., Bergelson J.M., Liddington R.C.;
 RT "Crystal structure of the I domain from integrin alpha2beta1.";
 RL J. Biol. Chem. 272:28512-28517 (1997).
 RN [4]
 RP VARIANT HPA-5 (BR).

RX MEDLINE=94043762; PubMed=7501236;
RA Santoso S., Kalb R., Walka M., Kiefl V., Mueller-Eckhardt C.,
RT Newman P.O.;
RT "The human platelet alloantigens Br(a) and Br(b) are associated with a
RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
RT alpha 2).";
RL J. Clin. Invest. 92:2427-2432 (1993).
RN [5]
RP VARIANT GLU-534.
RX MEDLINE=20206009; PubMed=10744142;
RA Kroll H., Gardemann A., Fechter A., Haberbosch W., Santoso S.;
RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
RT gene polymorphism on coronary artery disease and acute myocardial
RT infarction";
RL Thromb. Haemost. 83:392-396 (2000).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
CC associates with beta-1. Interacts with HPS5.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- POLYMORPHISM: POSITION 534 IS ASSOCIATED WITH PLATELET-SPECIFIC
CC ALLOANTIGEN HPA-5 (BR). HPA-5A/BR(A) HAS LYS-534 AND HPA-5B/BR(B)
CC HAS GLU-534. HPA-5B IS INVOLVED IN NEONATAL ALLOIMMUNE
CC THROMBOCYTOPENIA (NAT OR NATP). THE K534E POLYMORPHISM MAY PLAY A
CC ROLE IN CORONARY ARTERY DISEASE (CAD).
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17033; CAA34894.1; -;
DR EMBL; AF512556; AM34795.1; -;
DR PIR; A33998; A33998
DR PDB; 1A0X; 25-NOV-98.
DR PDB; 1DZ1; 02-AUG-01.
DR Genew; HGNC:6137; ITGA2.
DR MIM; 192974; -;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F:collagen binding activity; TAS.
DR GO; GO:0007596; P:blood coagulation; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Intergrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;

3D-structure. 1 29
KW SIGNAL 30 1181
FT CHAIN 30 1181
FT DOMAIN 1133 1154
FT TRANSMEM 1155 1181
FT DOMAIN 1155 1161
FT DOMAIN 1155 1161
FT REPEAT 45 103
FT REPEAT 7 2
FT DOMAIN 188 378
FT REPEAT 378 433
FT REPEAT 434 486
FT REPEAT 488 549
FT REPEAT 551 610
FT REPEAT 615 667
FT CA_BIND 499 507
FT CA_BIND 563 571
FT CA_BIND 627 635
FT SITE 1157 1161
FT DISULFID 83 92
FT DISULFID 680 737
FT DISULFID 789 795
FT DISULFID 865 876
FT DISULFID 1019 1050
FT DISULFID 1055 1060
FT CARBOHYD 105 105
FT CARBOHYD 112 112
FT CARBOHYD 343 343
FT CARBOHYD 432 432
FT CARBOHYD 460 460
FT CARBOHYD 475 475
FT CARBOHYD 699 699
FT CARBOHYD 1057 1057
FT CARBOHYD 1074 1074
FT CARBOHYD 1081 1081
FT VARIANT 534 534
FT TURN 170 171
FT STRAND 173 180
FT TURN 183 184
FT HELIX 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT STRAND 220 224
FT TURN 226 228
FT HELIX 232 240
FT TURN 241 241
FT HELIX 252 262
FT TURN 263 264
FT HELIX 266 268
FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT HELIX 292 301
FT TURN 302 303
FT STRAND 304 311
FT HELIX 313 317
FT TURN 318 319
FT HELIX 323 330
FT TURN 331 332
FT HELIX 337 340
FT STRAND 341 344
FT HELIX 347 353
FT HELIX 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;
Query Match 18.2%; Score 1069; DB 1; Length 1181;
Best Local Similarity 27.0%; Pred. No. 4e-64;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43;

dbSNP:1801106).
/FTid=VAR_003977.

GO: GO:0007160; P: cell-matrix adhesion; TAS.
 DR InterPro: IPR000413; Integrin_alpha.
 DR InterPro: IPR002035; WFA_A.
 DR Pfam: PF01839; FG-GAP; 3.
 DR Pfam: PF00092; vwa; 1.
 DR PRINTS: PRO1185; INTEGRINA.
 DR PRINTS: PRO0453; WVFADOMAIN.
 DR SMART: SM00191; Int_alpha; 4.
 DR SMART: SM00327; WFA_1.
 DR PROSITE: PS00442; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE: PS0234; WFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 1167 INTEGRIN ALPHA-10.
 FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1123 1145 POTENTIAL.
 FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 38 97 FG-GAP 1.
 FT REPEAT 2 2 FG-GAP 2.
 FT DOMAIN 167 350 WFA.
 FT REPEAT 365 427 FG-GAP 3.
 FT REPEAT 428 482 FG-GAP 4.
 FT REPEAT 483 545 FG-GAP 5.
 FT REPEAT 546 605 FG-GAP 6.
 FT REPEAT 608 660 FG-GAP 7.
 FT DOMAIN 1134 1140 POLY-LEU.
 FT CA_BIND 494 502 POTENTIAL.
 FT CA_BIND 558 566 POTENTIAL.
 FT CA_BIND 620 628 POTENTIAL.
 FT DISULFID 76 86 BY SIMILARITY.
 FT DISULFID 666 675 BY SIMILARITY.
 FT DISULFID 681 736 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 844 844 I -> L (IN REF. 2).
 FT CONFLICT 909 909 G -> V (IN REF. 2).
 FT CONFLICT 926 926 E -> D (IN REF. 2).
 SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;
 Query Match 18.1%; Score 1065.5; DB 1; Length 1167;
 Best Local Similarity 29.3%; Pred. No. 6.9e-64;
 Matches 358; Conservative 136; Mismatches 488; Indels 181; Gaps 46;
 QY 1 ENLDTENAMTQENARG-FQSGVVLQSGSR---VVVGAQEIIVAAQNRSLVQC-----50
 DB 23 FNLDEHPLRFPQPPAEFGYSVLQVHVGQQRWMLVGAWDGSGRRGDVTRCPVGGAH 82
 QY 51 -----DYSTG-SCSPIRLQVPVEAVNMSLGLSLAATSPQQLACGPTVHQTCE 99
 DB 83 NAPCAGHLGDYQLGNSHP-----AVNHLGMSLLETDGCGFMACAPLWSRACGS 134
 QY 100 NTYVKGCLFLFGNLRQOQFOKFPPEARLARGCPQEDSDIAFLDGSGLIIPHD-----PRRM- 153
 DB 135 SVFSGICARVDASFQPGQSLAPTAQR-CPTY-MDVVIVLDGNSIYPMWSEVQTFULRLV 192
 QY 154 -KEFVSTVWEOLKSKTSLSLMOYSEEFRIHFTFEFQNNPNRSLVKPITOLLGR-THT 211
 DB 193 GKLFIDP-EQIQ-----VCLVOYGSPVHEWSLGDGFRTEEVRRAKNLSRREGRETKT 245
 QY 212 ATGVRKVIREFLLNITNGARKNAFKILIVITDGSKF-GDPLGYEDVPEADREGVIRYIG 270
 DB 246 AQAIMVACTEGFSQSHGGRPEARLLVVVTDDGSHDEELPAALKACEAGR--VTRYGIA 303

RESULT 15

ITAL RAT

ID ITAL RAT

AC F18614;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

QY 271 V-GDAFRSEKS-----ROELNTIASKPPRDRHVQVNNFEALKTIQNLREKIFAIEGTQTG 325
 DB 304 VLGHYLRQRDPSSFLREIRTIASDPDERFFNFVNTDEAALTDIVDALGDRIFGLEGSHAE 363
 QY 326 SSSFEHEMSQEGFSAAITNSGPLLSVTSGVDWAGGVFLYTSKSEKSTFINNTRVDS----381
 DB 364 NESSFGLEMSQIGFSTHRLKDGLIFGMVGAYDMGCSVLWLEGGHRLPPRMALEDEFPFA 423
 QY 382 -DMNDAYLYA-AAIILNRVQSLVLCAPRYQHLGLVAMPR-ONTGMWESNANVKGTOIG 438
 DB 424 LQNHAAVLYGYSVSSMLLRGRRFLSGAPRFRHKGKVIAPOLKXGDAVRAVQAQGGBOIG 483
 QY 439 AYFASCLSDVDVDSNGSTDLVLI GAPHY--EOTRGQGVSVCPRLPRGQARQWQDAVLVG 496
 DB 484 SYFGSELCLPTDRDGTITDVLVAAPMFLGPQNKETGRVYVYLV-CQOSLTLTQGLQP 541
 QY 497 EQGQPMGRFGAALTVLGVNGDKLUTVAIGAPEGEDNRGAVYLFHGTSGSGSIFSHSORI 556
 DB 542 EPPQD-ARFGFAMGALPDNLQDGFADVAVGAPLEDHGQALYLYHGTQ-SGVRPHPAQRI 599
 QY 557 AGSKLSPLQVYGOSLSGGQDLTWDGLVLTVAAGHVLLLRSSOPVLRVKAIMENPREV 616
 DB 600 AAASMPHALSYFGRSVDRGLDLDGDDLVAVAGAAIILSSRPVHLTSPSLEVTPOAI 659
 QY 617 ARNVFECNDQVVKGEAG--EVRVCLHVQKSTRDLREGQIQSVVTVYDLALDSGRPHSRA 674
 DB 660 SVVQRDCRR---RQGEAVCLTAALCFQVTSRTPGRWDH--QFYMRFTASLDEMTAGARA 713
 QY 675 VFNET--KNSTRQTVLGLTQCTETLKLQIKNCIEDPVSPVILRLNPLVGTPLUSARGN 732
 DB 714 AFDGSGQRLSPRLRLSVG-NVTCEQLHFHVLDTSDYLRPVALVTTPALDNTTKPG---768
 QY 733 LRPVLAEDAQLFTALPFEKNCGNDCNLCQDDLSITFSFMSLDC-----LVVGGPRE 784
 DB 769 --PVLNEGSPTSIOKLFPFKDCDPCNECVTDVLQ---VMDIRGSRKAPFVVVRGRRK 823
 QY 785 FNVTVVRNDCGDSYRTQVTFPFDLDSYRKVSTLQNRORSWRLACESASSTEVSGAL 844
 DB 824 VLVSTTLNENKENVNTSLSIIFSRNL---HLASLTPQR-ESPIKVECAAPSA-----872
 QY 845 KSTCSINHPITPENSEVTFNITFDVDSKASIG---NKLL-----LKANVTSENMPRTN 896
 DB 873 HARLCSVGHVPFQTAGKVTFLLEFEFSCSLLSQVFGKLTASSDLSLRNGTLQENTAQT-931
 QY 897 KTEFOLEPLVKYAVYVMTVSHGVSTKYNLFTASENTSRVMQHY-----Q 941
 DB 932 -----SAYIQEPH-----LLFSESTLHRYEHPYGTLPVGPGEPEKPTTLR 973
 QY 942 VSNLG---QRLSLPISLVFLVP-----VRLNQTIVDRPQVTFSENLSSTCHTKERL 989
 DB 974 VQNLGCVVVGSLIISA--LLPAVAHGNVYFLSLQVI-----TNVASCIVQNLTEP 1022
 QY 990 PSHSDFLAELRKAPVNVNCIAVCQRIQCDIPFFGQEEFNATLKNLSDFWYIKTSHNHL 1049
 DB 1023 PGPPVHPBELQHTNRLNGSNTQCQVVRCHLGAKGTEVSGVLLRLVHNEFPFRKFKSL 1082
 QY 1050 LIVSTABELFNDSTVFLPGQGAQVRSQTEKVEFEVNPPLPLIVGSSVGLLILALIT 1109
 DB 1083 TVVSTFELTEBGSVLQLTEASRWSESLLEV-VQTRPILISLWILIGSVLGLLLALLV 1141
 QY 1110 AALYKLGFF-----KROYK 1123
 DB 1142 FCLWKLGFPAHKKIPEEKREK 1164

STANDARD; PRT; 1180 AA.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 25, 2003, 14:11:54 ; Search time 33.1068 Seconds
(without alignments)
8862.397 Million cell updates/sec

Title: US-09-902-481b-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARFGQ.....FKRQYKDMWSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rviro.*
16: sp_bacteriap.*
17: sp_archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length DB ID	Description	
1	4387	74.7	1151 11	Q9J130	Q9J130 rattus norv
2	3907.5	66.5	1036 11	Q8CA73	Q8CA73 mus musculus
3	3815.5	64.9	920 6	Q28984	Q28984 sus scrofa
4	3484	59.3	1169 4	Q8IVA6	Q8IVA6 homo sapien
5	3326.5	56.6	1169 11	Q9QXH4	Q9QXH4 mus musculus
6	3251.5	55.3	1161 11	Q9QYE7	Q9QYE7 rattus norv
7	1529.5	26.0	1161 11	Q9MTV4	Q9MTV4 mus musculus
8	1519	25.9	1160 11	Q9R200	Q9R200 mus musculus
9	1414	24.1	1196 13	Q98TF1	Q98TF1 cyprinus ca
10	1363.5	23.2	1187 13	Q98TF0	Q98TF0 cyprinus ca
11	1350.5	23.0	1086 4	Q96HB1	Q96HB1 homo sapien
12	1270	21.6	927 6	Q8HZV0	Q8HZV0 bos taurus
13	1163.5	19.8	1167 11	Q88340	Q88340 rattus norv
14	1119	19.0	1167 11	Q88341	Q88341 rattus norv
15	1051	17.9	1171 13	Q42094	Q42094 gallus gall
16	1041	17.7	1038 11	Q8BS01	Q8BS01 mus musculus

17	1014	17.3	895 11	Q9WUF8	Q9WUF8 mus sp. itg
18	1013.5	17.3	1160 6	Q8MKF4	Q8MKF4 felis silve
19	865	14.7	348 4	Q8TES5	Q8TES5 homo sapien
20	852	14.5	1332 5	Q8BPQ8	Q8BPQ8 halocynthia
21	807	13.7	205 11	Q83001	Q83001 rattus norv
22	753.5	12.8	780 13	Q06271	Q06271 xenopus lae
23	738	12.6	823 4	Q8WY18	Q8WY18 homo sapien
24	686.5	11.7	823 11	Q8CE84	Q8CE84 mus musculus
25	669	11.4	1032 11	Q61989	Q61989 mus musculus
26	643	10.9	1036 11	Q91YD5	Q91YD5 mus musculus
27	640.5	10.9	1033 6	Q9BGU3	Q9BGU3 bos taurus
28	630.5	10.7	257 11	Q8C270	Q8C270 mus musculus
29	619.5	10.5	1041 5	Q9UB90	Q9UB90 lytechinus
30	614.5	10.5	1041 5	Q76378	Q76378 lytechinus
31	579.5	9.9	1054 5	Q9U6S1	Q9U6S1 strongyloce
32	579	9.9	1034 13	Q98TT7	Q98TT7 gallus gall
33	550	9.4	1033 13	Q42598	Q42598 xenopus lae
34	534	9.1	1016 13	Q91779	Q91779 xenopus lae
35	530	9.0	974 11	Q924W2	Q924W2 rattus norv
36	529	9.0	1073 11	Q8CC06	Q8CC06 mus musculus
37	526	9.0	1047 6	Q9MZD6	Q9MZD6 bos taurus
38	525.5	8.9	1007 6	Q9GK48	Q9GK48 bos taurus
39	512.5	8.7	1034 6	Q9TUN4	Q9TUN4 oryctolagus
40	509.5	8.7	1036 6	Q9TUN6	Q9TUN6 sus scrofa
41	506.5	8.6	1012 11	Q70304	Q70304 mus musculus
42	491.5	8.4	1049 5	Q8SY51	Q8SY51 drosophila
43	475	8.1	833 5	Q9BPQ7	Q9BPQ7 halocynthia
44	469.5	8.0	1036 6	Q9TU44	Q9TU44 canis famil
45	468.5	8.0	604 11	Q8BQ25	Q8BQ25 mus musculus

ALIGNMENTS

RESULT 1

Q9J130
ID Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerrila K. Jr.;
RT "Cloning of the rat CD11b cDNA sequence."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268593; AAF81280.1; -
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5_
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWF; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.7%; Score 4387; DB 11; Length 1151;
Best Local Similarity 73.1%; Pred. No. 3e-313;
Matches 831; Conservative 146; Mismatches 158; Indels 2; Gaps 2;
Qy 1 FNLDTENAMTFQENARFGQSVVQLQGETRVVVAAPQEVKAVNOTGALYQCDYSTNRCPI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQGETRVVVAAPQEVKAVNOTGALYQCDYSTNRCPI 76


```
Db 259 TDRGKQDNLSYSDVIPMAEAAIIIRYAIAGVKAIFYNEHKSQBLKAIASMPSEHYVFSVE 318
Qy 301 NFBAKTIQNLQREKIPAIETGTTGSSSSFEHEMSQEGFSAATISNGPLLSITVGSVDWAG 360
Db 319 NFDAKDIEQLQEKIPAIETGTTGSSSTFELEMSQEGFSAVTPDGVVUGAVGSFWSG 378
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAAAIILNRNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 379 GAFLYPSNMPTTFINMSQENMDRDAYLGYSTALAFWKGVHSLILGAPRHOHTGKVIFT 438
Qy 421 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGQGVSVCLP 480
Db 439 QESRHRPKSEVRGTQIGSYFGASLCSVDMDRDSGTDVLIGVPHYEYTRGQGVSCPM 498
Qy 481 PRGORARWQCDVLYGSGQWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 499 P-GVGSRWHCGTTLHGSGQHPWRFGAALTVDGVNGDSLADVAIGAPGEENRGAVYIF 557
Qy 541 HGTSGSGISPSHRSORIASGSLKSLPRLOYFGQSLGGQDLTMDGLVDLTVGAGQHVLRLSQ 600
Db 558 HGASRQDIAPSPORISASQIPSRIOYFGQSLGGQDLTMDGLVDLAVGSKGRVLLRTR 617
Qy 601 PVLRVKAIMBFNPREVARNVFPCNDQVVKGEAGEVRVCLVHVKSTRDRIRREGIOISVVT 660
Db 618 PILRVSPVHTFAEISRSVFECQEQVAPEQTLSDATVCLHIESPKTQL--GDLRSTVT 675
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLGTCTCETLKLQLPNCIEDPVSIVLRLNF 720
Db 676 FDUALDHGRULSTRAIFKETKTRALTRVKTGLGNKHCESVKLLPACVEDSVTPIURLNF 735
Qy 721 SLVGTPLSAFCNLRPLVAEDAQRALTALFPFEKNCNDNIQQDLSLTFPSFMSLDCLVVG 780
Db 736 SLVGVPISSLQNLQPLAVDDQTYFTASLPFKNCNCGADHICQDDLSVVFDPDLKTLVVG 795
Qy 781 GPREFNVTVTRNDGDSYRTQVTFPPFLDLVRYKYSTLQ-----NQRQSRWR 829
Db 796 SDLELNDVTVSNDGDSYGTVTFLFPVGLSFRRAEGQVFLRKKBQDQWQRGQSHL 855
Qy 830 LACESASSTVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSE 889
Db 856 LMCD--STPDRSQGLWSTGSSRHVIFRGGSQMTFLVTFDVS PKAELGDLRLLRARVGE 913
Qy 890 NNPRTNKTBFQLELPVKYAVYVMTVSHGVSTYKLNFTASE--NTSRVMQHYQVSNLQOR 948
Db 914 NNVPCTPKTTFQLELPVKYAVYVMTVSHGDQFTKYNLNFSTSEKETSVEHRFQVNNLQOR 973
Qy 949 SLFISLVLPVRLNQTIVDRPQVTFSENLSSTCHTKERLPKSHDPLAELRKAPVNVCS 1008
Db 974 DVPVSINFVPIELKGEAVN--TVMVSHPNPLTQCYRNRLKPTQFDLLTHMQKSPVLDCS 1032
Qy 1009 IAVQRIQCQDIPFGIOEEFNATLKNLSFDWIKTSHNHLILVSTABILFNDSVFTLLP 1068
Db 1033 IADCLHLRCDIPSLGIULDELYFLKGNLSFGWISQTLQKVLGLSEAEIETNTSVYSQLP 1092
Qy 1069 GQGAFRVRSQETKVEPPEVENPLIIVGSSVGLLALLALITAAIYKLGFFKRYOKDMNSE 1128
Db 1093 GQBAFLRAQTKVLEMTKRNPNPLIIVGSSVGLLALLALITAIYKAGFFRQTKEMLEE 1152

RESULT 6
Q0QY67 PRELIMINARY; PRT; 1161 AA.
AC Q0QY67;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAP21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR HSSP; IPR002035; VWF A.
DR PFam; PF01839; FG-GAP; 5.
DR PFam; PF00357; integrin_A; 1.
DR PFam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWF A; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.3%; Score 3251.5; DB 11; Length 1161;
Best Local Similarity 58.0%; Pred. No. 9.8e-230;
Matches 654; Conservative 158; Mismatches 302; Indels 13; Gaps 9;

Qy 2 NLDTENAMTFOENARFGQSVVQLQGSRVVVGAPQEIIVAAANORGLYQCDYSTGSCPIR 61
Db 21 NLDVEPIVFREDASFGQTVVQFGSRLVVGAPLEAVNQTGRLYDCAPATGMQPIV 80
Qy 62 LQVPEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGLCFLFGSNLRQPPQKF 121
Db 81 LRSPLAEVNSLGLSLVATNNAQLLACGPTAQACRVAKVAKGSCLLGSSL-QFTQAV 139
Qy 122 PEALRCPCPEDSIAFLIDGSGSIIIPHDFRMKEFVSTWMEQLKSKTILFSLMOYSEFR 181
Db 140 PASMPCPCQEMDIAFLIDGSGSINQDFQAKMDFYKALMGEPASTTLFSLMOYSNLK 199
Qy 182 IHFTFEFQNNPNRPLVKPIITOLGRTHRTATGRVKVIRELLNITNGARKNAFKILVIT 241
Db 200 THFTTFEFKNLDLPQSLVDPIVOLQGLTTATGIRTVMEELFHSKNGSRKSAKKILLVIT 259
Qy 242 DGEKFGDPLGYEDVIEPADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDHVFQVN 301
Db 260 DGQYRDPLEYSVDIIPADKAGIIRVIGVDGAFQEPALKELTNTIGSAPPQDHVFKVN 319
Qy 302 FEALKTIONLREKIPAIETGTTGSSSSFEHEMSQEGFSAATISNGPLLSITVGSYDWAG 361
Db 320 FAALRSIQRLQEKIPAIETGTTGSSSSFEHEMSQEGFSAATISNGPLLSITVGSYDWAG 379
Qy 362 VFLYTSKESKSTFINMTRVDSMDNDAYLGVAAAIILNRNRVQSLVGLGAPRYQHIGLVAMFR 421
Db 380 AFLYPNTPTTFINMSQENMDRDAYLGYSTALAFWKGVHSLILGAPRHOHTGKVIFTQ 439
Qy 422 NTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGQGVSVCLP 481
Db 440 EABHWPKESEVRGTQIGSYFGASLCSVDVDRDGSITDLVIGAPHYYEQTRGQGVSVFPV 499
Qy 482 RGRARWQCDVLYGSGQWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 541
Db 500 -GVRGRWQCEATLHGSGQHPWRFGAALTVDGVNGDNLDVAIGAPGEESRGAVYIF 558
Qy 542 GTSGSGISPSHRSORIASGSLKSLPRLOYFGQSLGGQDLTMDGLVDLTVGAGQHVLRLSQ 601
Db 559 GASRLIMPSPSRVTVGSQSLRLQYFGQSLGGQDLTQDGLVDLAVGAGQHVLRLSLP 618
Qy 602 VLRVKAIMBFNPREVARNVFPCNDQVVKGEAGEVRVCLVHVKSTRDRIRREGIOISVVT 661
Db 619 LLKVELSIRFAPMEVAKAVYQCHERTPTVLEAGEATVCLTVHKGSPDLL--GNVQGSVRY 676
Qy 662 DLALDSGRPHSRVAFNETKNSRTRQVGLGTCTCETLKLQLPNCIEDPVSIVLRLNFS 721
Db 677 DLALDPPGRULISRAIFDETKNCTLTGRKTLGLGDHCETVKLLPDCVEDAVSPIILRLNFS 736
```

QY 722 LVGTPLSAFNGLRNPLVLAEDAQRLFTALPFFPKNCGNDNICODDLSITSPMSLDCLVVGG 781
DB 737 LVKDSASP-RNLHPVLAVGSDPHITASLPFFKCKCKQELLCEGLGIFSNFGLQVLVVG 795
QY 782 PREFNVTIVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRORSORSMRLACESASSTVS 841
DB 796 SPFLVTVTVNNEGDSVGLVKFYYPAGLSYRRTGQ-QPHQYPLRLACEAPAAQED 854
QY 842 GALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENPRNKTTFQ 901
DB 855 --LRSSSCSINHPIPREGAKTTFMITFDVSKAFGLDRLLRKASSSENKPDNTKTAFA 912
QY 902 LELPKYAVYVTVNNEGDSVGLVKFYYPAGLSYRRTGQ-QPHQYPLRLACEAPAAQED 960
DB 913 LELPKYAVYVTVNNEGDSVGLVKFYYPAGLSYRRTGQ-QPHQYPLRLACEAPAAQED 972
QY 961 RLNQTVIVDRPOVTFSENLS--TCHTKERLPSHSDFLAELRKAPVYVNGSIACVQRIQCDI 1019
DB 973 LINGVAVMD--VTLSPPAQGVSCVQMKPPQNPDPFTQIQRSVLDCSIADCLHFRCDI 1029
QY 1020 PFGIOEENATLKNLGFWDYIKTSHNHLIVTAEILFNDSVFTLLPGQAFVRSOTE 1079
DB 1030 PSLDIQDELDFLGNLSFGWVSQTLQEKVLLVSEAITFDTSVYSQPLPGQEAFLRAQVE 1089
QY 1080 TKVEPFEVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKDMW 1126
DB 1090 TTEEBYVYVEPIFLVAGSSVGLLILALITAAALYKLGFFKQYKDMW 1136

RESULT 7

Q9WTV4 PRELIMINARY; PRT; 1161 AA.
ID Q9WTV4;
AC Q9WTV4;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Spleen;
RA Ma R.Z.; Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; RAD25885.1;
DR HSSP; P20701; ILFA.
DR MGD; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWF_A; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;

Query Match 26.0%; Score 1529.5; DB 11; Length 1161;
Best Local Similarity 33.8%; Pred. No. 4.3e-103;
Matches 396; Conservative 220; Mismatches 456; Indels 101; Gaps 37;

QY 1 FNLDTENAMTFQENA-RFGQSVVQLQGSRRVVGAPQEIIVAAQNRGLSYQCDYSTGSCBP 59
DB 24 YNLDRTPQSLAQACRHFGYQVLQIEDG-VVVGAPGE--GDNTGGLYHCRSTSEFCBP 79
QY 60 IRLQVPEAVNMSLGLSLAATSPPOLLAGPTVHQTCSNTYVYKGLCFGLFSGNLRQOPQ 119

DB 80 VSLH-GSNHTSKYLGWTLATDAAGKSLACDPLGSRTCQNTYLSGLCYLFPQSLGPM 138
QY 120 KFEALRGCEQSDSDIAFLIDGSGSIIPHDFRRKKEPVSTVMEQLKSKTKLFSLMQVSEE 179
DB 139 QNRPAYQECMKGVLDLFLDFGSGSLDRKQEKILEFMKQVMRKLSNTSQFAAVQFSTD 198
QY 180 FRIHFTPEF-QNNPNRSLVKPITOLLGRTHATGVKRIRELLNITNGARKNAFKILI 238
DB 199 CRTEFTFLDVQKNKNDVLLGSVQPMFLTNTRAINVYVAHVKEESGARPDATKVLV 258
QY 239 VITDGEKFGDPLGYEDVPEADREG-----VIRYVIGVDADFRRSEKSOELNTIASKP 291
DB 259 IITDG-----EASDKGNISAAHDITRYIIGIGKHFVSQVKQKTLHIPASEP 304
QY 292 PRDHVQVNNFEALKTTQNLQREKIPAIETGTQSSSSFEHMSQEGFSAAITNSGPLLS 351
DB 305 VEEFVKILTDFEKLKOLFQDQRRYIAIEGTRQDLTSFNMELSSSGISADLSKGHAVG 364
QY 352 TVGSYDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGVAAA-IILNRVQSLVLGAPR 409
DB 365 AVGAKWAGFLLDREDLOGATFVGQPLTSDVRCGYLGTVAWMTSRSRPPLAAGAPR 424
QY 410 YQHIGLVAMPR--QNTGHWESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYY 467
DB 425 YQHVQVLLFQAPPEAGRWNTQKIEGTQIGSYFGGELCSVDLDQDGEAELLIGAPLFF 484
QY 468 EQTRGGQVVCPLPRGORARWQCDALYBGOQPWGRFGAALTVLGVDVNGDKLTDVAIGA 527
DB 485 GEQGRGVFTY--QRRQSUFEMVSELQDGPYPLGRFGAALTALTDINGDRLTDVAVGA 541
QY 528 PGEEDNRGAVYLFHGTSGSGISPSHSORIASGKLSPRLOFGQSLSGGODLTMDGLVDLT 587
DB 542 PLEE--QCAVYIFNGKFG-GLSPQPSQRIQAGVDFGIRWFGRSIHGVDLGGDRLDV 598
QY 588 VQAQGHVLLLRQPLRVKAIMFENPREARNVFCNDQVVKGEAG-EVRVCLRHVKST 646
DB 599 VGPEGRVVVLSRRPVDVVTLSFSPREIPVHEVECSVSAREEQKHGKVLKACFRKPLT 658
QY 647 RDRUREGOIQSVTYDIALDSGRPHSRVAFNETKNSTRQTVGLGLTQCTETKLQLPNC 706
DB 659 PQ--FOGRLLANLTYTLQDHRMRSGLPDGSHELSGNITSITP-DKSCLDHFHFHFPIC 715
QY 707 IEDPVSPVILRLNFSLV--GTPLSAFGN-LRPLVAEDAQRLFTALPFFKNCNDNICQ 762
DB 716 IQDLISPINVSLNLSLEEETPRDQGRAMQPILRPSIHTV-TKEIPFEKNCGEDKCE 774
QY 763 DDLSTTFMSLDCLVVGGP-----REFNVTVTRNDGDSYRTQVTFPPDLDSYRKV 816
DB 775 ANLTLSPPARS-----GPLRLMSSASLAVETLNSGSDAYVWRLDLPFRGLSPFRK 827
QY 817 STLQNRQSRQSWRLACSSASSTEVSGAL-KSTSCSINHPIPPENSEVTNITFDVDSKAS 875
DB 828 EMLQ---PHSRMPVSCEEL--TEGSSLLTKLKNVSSSPIFKAGQEVSLQVMTLLNSS 882
QY 876 LGNKLILLKANVTSEN-NMPTNKTFQLEPLVKYAVYVTVNTHSGVSTKYLAFNTASEN 934
DB 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLPVNLITKEQENSTLYISFTPKGPKIQ 942
QY 935 VMQHOYQV--SNLQQRSLPISLFLVPLVRLNQTIVMDRPO--VTFSENLS--TCHT 985
DB 943 QVQHVYQVRIQPSAYDNMP-TLEALVGVP-----WPHSEDPITYTWSVQTDPLVTC 995
QY 986 KE-RLPSSHSDFLAELRKAPVYVNGSIACVQRIQCDIIPFGIQEEFNATLKNLSFDWYIKT 1044
DB 996 EDLKRPSSE--AEQPCLPGV-----QFRCPITF---RREILIQVTGTVELSKELKA 1041
QY 1045 SHNHLIIVSTAEILLFNDVSFTLLPCOGAFVRSOTETKVEFEVNPPLPLIVGSSVGGILL 1104
DB 1042 S-STLSLCSLSVSFNSKHFHLYGSKA-SEAQVLVVDLITHEKMLHVYVLSGIGGLVL 1099
QY 1105 LALITAAALYKLGFFKQYKDMW-SEGGPPGAEP 1136

```
Db 1100 LFLIFLALYKVGFFKRNLEKMEADGGVNGSP 1132
RESULT 8
Q9R200 PRELIMINARY; PRT; 1160 AA.
AC Q9R200
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065901; AAD25884.1; -
DR HSP; P20701; 1LFA.
DR MGB; MGI:96606; Itgal.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 5
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA_1
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1160 AA; 128127 MW; A33C531B139F1FAD CRC64;

Query Match 25.9%; Score 1519; DB 11; Length 1160;
Best Local Similarity 33.6%; Pred No. 2.5e-102;
Matches 394; Conservative 219; Mismatches 459; Indels 100; Gaps 36;

Qy 1 FNLDTENAMTFOENA-RFGGOSVVOQGSRVVGAPQBIIVAAVQKSLYQCDYSTGSCPE 59
Db 24 YNLDRPTQSLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRSTSEFCQP 79
Qy 60 IRLQVPVEAVNMSGLSLAATTSPPQLACGPTVHOTCSNTYVYKGLCFGLGSLNLRQOPQ 119
Db 80 VSLH-GSNHTSKYLGMTLTADAAGKSLACDPGLSRTCDQNTYLSGLCYLPQSLGPM 138
Qy 120 KFPALRGCPQEDSDIAFLIDGSGSII PHDFRRMKFEFVTVMEQLKSKSLTFLSLMOYSEE 179
Db 139 QNRPAYQCMKGVLDLVEFDSQSLDRKDFEKLFEKMDVNRKLSNTSYQFAAVQFSTD 198
Qy 180 FRIHFTFKPE-QNNPNRSLVKPIITQLGRTHATGVRKVIARELLNITNGARKNAFKILI 238
Db 199 CRTEFTFDYVQKKNPDLVLLSGVQPMELLNTFPAINVVAHVFKESGARPDPATKVLV 258
Qy 239 VITDGEKFGDPLGYEDVTEADREG-----VIRYIVGVGDAPRSEKSRQELNTIAKSP 291
Db 259 IITDG-----EASDKGNISAAHDITRYIIGIKGHFVSVQKQLTHIFASEP 304
Qy 292 PRDHVFOVNNFEALKTIONLREKIFAIEGTQTGSSSFEHEMSQEGFSAITSNGLPLLS 351
Db 305 VEEFYKIILDTFEKLKDLFTDQRRYYAIEGNRQDLTSFNMELSSSGISADLSKGHAVVG 364
Qy 352 TVGSDVWAGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAA-IILNRVQSLVGLGAPR 409
Db 365 AVGAKDWAGGDLREDLQATFVQCEPLTSDVRGCVGLGYTVAVMTSSRSPPLAAGAPR 424
Qy 410 YQHIGLVAMFR--QNTGWMESNANVKGTOIGAYFGASLCSDVDSDNGSTDLVLGAPHY 467
Db 425 YOHVQVLLFQAPAEAGGRWNTOKIEGTIGSYFGGLCSVDLDQDGEALLIGAPLFF 484
```


DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 5.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR PRINTS: PRO0453; VWFADOMAIN.
DR SMART: SM00191; Int. alpha; 5.
DR SMART: SM00327; VWA_1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0234; VWF_A; 1.
SQ SEQUENCE 1196 AA; 132477 MW; 9369C807E7DCA53B CRC64;

Query Match 24.1%; Score 1414; DB 13; Length 1196;
Best Local Similarity 31.3%; Pred. No. 1.4e-94;
Matches 373; Conservative 227; Mismatches 446; Indels 144; Gaps 39;

QY 1 FNLDTENAMTFOENARG-FGQSVVQLQ-GSR--VVVGAPOEIVAAQNRGSLYQCDSYSGS 56
DB 32 FNIDTEHPLRFNGAPEDFFGYSVYQTEFGNRKQIIVGAPLE---GNSGEMYSCTADLQGS 88
QY 57 CEPILRQVP-----VEAVNMSLGLSLAATSPQLLACGPTVHQTSENTRYKGLCFLFGS 112
DB 89 CQ--RLQRPGESEVRFVFGMSAAVSSAALTS-----CSGYFPHECDGNSYLVNGVCYQFSS 140
QY 113 NLRQOPQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOLKSKTLFS 172
DB 141 SL-QAVSNFTAYQCSKREVNVLVDFGSSSMKTVDFEMKNKVFIDIMKKLSNISKFA 199
QY 173 LMOYSEERIHFTFKFQNNPNRSLVKPITQLLGRTHATQVRKVIKRELN-INGARK 231
DB 200 AVQFSDVTRTFDFNDYQSGSAEKLKMBE-THMKSILTNTHKAIDVILKNLLNSMLSGADS 258
QY 232 NAFKILIVTDGKFGDPLGYED--VIPADREGVIRYVIGVDAPFRSEKSRQELNTIAS 289
DB 259 KAQKALVIITD---GDPSDNDYVNLKCDQNILYIIVG-----KVDILIELTQLAS 309
QY 290 KPRPDHVFQVNFPEALKTQONLREKI PAIEGTQTGSSSSFEHMSQEGFSAATNSGFL 349
DB 310 EPRRNTTFIKYDGLKGLDNLQKLYNIESVDAQGRDKQELKSQSGFSVYQEEVSI 369
QY 350 LSTVGSYDWAGGVFLYTSKESTFINMTRVDSMN-DAYLGYAAAIILNRVQSLVLGAP 408
DB 370 VGSVGSNDWRGALYEVTG-SGSDFKETETIDPAVNKDSYMGYSTVLGMRGVSLLFSGAP 428
QY 409 RVQHIGLVAMFQNTQWGESNANVKGTQIGAYFGASLCSVDVDSNSTDLVLIGAPHYVE 468
DB 429 RAETGLVTLFTKNQNTWTVTNSINGEQIGSYFGASLSLDDVSDGSDFLVLGAPLFYQ 488
QY 469 QTRGGQVSCPLPRGORARWQCDVAVLYGEO-----QPMGRFGAALTVLGDVNGDKLTDV 523
DB 489 SQ-----PRTEGRLYVYSLSEQYFQKTLNVQSSTTGRFAASVASLKDLNGDGLSDV 540
QY 524 AIGAPGEEDNRGAVYLFHTGTSIGSIPSHS-QRIAGSKLSPLRQVFGQSLSGQDLTMWG 582
DB 541 AVGAPLE--NEGYYVYVLGDRTHGINPELTQORISVQSLPLGQVGLQPGVSLTQGMNDNDN 598
QY 583 LVDLTVGAQHVLLRSQVLRVKAIMEPNPREVARNVECDNDQVYKGEAGEVRVCLHV 642
DB 599 LTDIVIGAOGGIVLLKARPVMSYSAQLSPKSEISLNYFECPS--NAPNANLNTSCFTV 656
QY 643 QKSTRDLREGQIQSV--VTYDLALDGRPHSRVAFNETKNSTR--RQTVLGLTLTCTBT 698
DB 657 TERTSS---TGSLEKLNLSNLNVDRVGRMSRGDFDQSVSRITLQQSVLLDSSGSCFN 713
QY 699 LKQLPNCIEDPVSPILVRLNLSVGTPLSAFGLNLRPLVLAEDAQRLFTALFPFKNCNGND 758
DB 714 ESIFMLRCVADTVSPILKIRNFS--QTEMLS-GNSVAVLDVHSRTEENVEVPFQRCNSN 770
QY 759 NICQDILSTTFSPMSLDCLVWGPREFNTVTVRNDGEDSYRTQVTFPDLISYKQVST 818
DB 771 NSCVADLKLNFSTN-NTLVENQARHFTVQVSLANFCDSDSYNTSIVLHYPEGISLSKFDA 829
QY 819 LQNRQORSWRLACESASSTEVSGALKSTSCSINHPIFFENSEVTFNITFDV---DSKAS 875

DB 830 IKPSRTR-----SSCGDRDGSATNRTTCSIDLPPYRSGTTTQFLGTRFVVKWWDSS 881
QY 876 LGNKLLLLKANVTSNNMPTNKTEFOLELPVKAVVAVVVTSHGV-STKYLNFTASENSTR 934
DB 882 NRMBEIMITANSNNNM---SDTEVRRSVPVQFAVDLAISLVAEDSVTYMNFSLSDRGPK 938
QY 935 VMQHQYQVSNLQGRSLPISLVFLVPVR-----LN 963
DB 939 PLNTITYKVENSGRKLDPVSVTLTLPCTPHVLLTPHTFSMHEVHHSFISSVHQIIMCLLN 998
QY 964 QTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQ-----C 1017
DB 999 KHLFFSPSLSAVQVQVSTGSLREVC-SQF--DLNKSSAVHFNLTADARLQNVKEYES 1055
QY 1018 DIPFGIQEE--FNATLKGNSLFDWYIKTSHNHLIIVSTAEILLFNDVSVFTLLPGQGAFAVR 1075
DB 1056 KYSEYFERKONVFSISAEIINTNTSLYNQTS-----SELKYNPH-----R 1094
QY 1076 SQTETKVEPPEVNPPLIV-GSSVGGLLLLIALLIATAALYKLGFFKQRYKD 1124
DB 1095 SQTETKVE-FVVPFSLMLIVCTGAVGGFFLLIILFLLKCGFFKRNRPD 1143

RESULT 10
Q98TF0 PRELIMINARY; PRT: 1187 AA.
AC Q98TF0;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE C11-2.
GN CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB048537; BAB39135.1; --
DR HSSP: P20701; 1LFA
DR InterPro: IPR001969; Aspartatease_site.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 5.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR PRINTS: PRO0453; VWFADOMAIN.
DR SMART: SM00191; Int. alpha; 5.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0234; VWF_A; 1.
SQ SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;

Query Match 23.2%; Score 1363.5; DB 13; Length 1187;
Best Local Similarity 31.5%; Pred. No. 7.2e-91;
Matches 375; Conservative 209; Mismatches 451; Indels 157; Gaps 43;

QY 1 FNLDTENAMTFOENARG-FGQSVVQLQ-GSR--VVVGAPOEIVAAQNRGSLYQCDSYSGS 56
DB 32 FNIDTEHPLRFNGTPEDFGYSVYQTEFGNRKQIIVGAPLE---GNSAGEMYSCTADLQGS 88
QY 57 CEPILRQVP-----VEAVNMSLGLSLAATSPQLLACGPTVHQTSENTRYKGLCFLFGS 112
DB 89 CK--RLQRPGESEVRFVFGMSAAVSSAALTS-----CSGYFAHEDGNSYLVNGVCYQFNS 140
QY 113 NLRQOPQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOLKSKTLFS 172

Db 141 SL-QAVSNFTAAVQECSEKREVNLFVDFDGSSEKMAVEFDMNKNFIKDWKMLSNSSIKFA 199
Qy 173 LMOYSEEFRIHPTFKFQNNPNRSLVKPITQLGRTHATGVRKVIKRELLN-ITNGARK 231
Db 200 AVGFSTIRTVDFDNTQNSAEKLMKE-RHKSLTNTYKAINVLDKVNLSVSSGADP 258
Qy 232 NAFKILVITDGEKFGDPLGYED--VPEADREGVIRYVIGVGDAPFSEKSRQELNFIAS 289
Db 259 NAQKALVIITD---GDPSNDNDYILNICEQNILAYIIGV---KVDLTTLTQLAA 309
Qy 290 KPRDRHVQVNNFEALKTIONLREKIFATEGTGTGSSSSFEHMSQEGFSAITSNGPL 349
Db 310 EPLKNTFTYIQEYNGGLDLNLOKTYINIEGSEKARQKELSQSGFVVYQESVI 369
Qy 350 LSTVGSYDAGGVFLYT---SKEKSTFINMTRVDSMDNDAYLVGAIAIILNRNVQSILVLG 406
Db 370 VGSVGSNDWALYEVNGSGSKFQTEITDPAVN---KDSYMGSTVILGMRHGVSLFLSG 426
Qy 407 APRYOHTGLVAMFRONTGMHESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHY 466
Db 427 APRAEHTGLVTLFTKNSTVTVNRNINGBQIGSYFGASLSLLDVSQDSDFLVLGAPLF 486
Qy 467 YE-QTRG-QGVSVCLPRGRARQWQDAVLYGEO--QOPWGRFGAALTVLGVNVDKLTLD 522
Db 487 YQSOPRAEGRLYVYTL-----SEQYSQKTLQSTTGREFATSLASLKDLNGDGLSD 535
Qy 523 VATGAPCEEDNRGAVILFPHGTSGSGISPSHS-QRIAGSKLSPRLQYFGOSLSGGQDLTMD 581
Db 536 VAVGAPLE--NEGYYVITYLGDGTHGINPEHAPQIPARSVLPLGLQFGVGSLSQMDMND 593
Qy 582 GLVDLTGAAQHVLLLRQVLRKAIWENPREVARNVFECNDQVVKGEAGEVRVCLH 641
Db 594 NLDPDIVGTGGIVLLNARPMVSVAQSLSNPHEISLNYFECPGS--NAPNANFLTSCFT 651
Qy 642 VQKSTRDLREGQIQSV--VYDIALDSGRPHSAFNETKNSTR--RQTVLGLTQTC 697
Db 652 VTERTS--TGSLEKKLVSLNLDVVRGMSRGFPDMPDSSSRTLQOQSVLLDSGSSCS 708
Qy 698 TLKQLPNCIEDPVSPIVLRNFS---LVGTPLSARGNLRPVLAEDAQRLFTALFFFX 753
Db 709 NFSIFMLRCVADTVSPLKIRNMFSTQMLSGNSLAVL-DIQSRTEEYEVVL-----FQR 761
Qy 754 NCGNDNICQDLSITFSFMSLDCLVGGPREFNVTVTRNDGEDSVYRTVTFEFLDLSY 813
Db 762 NC-NSNSCVADLKLNFSTN-DTLVENVQAHFTVLVSLANPGDSDSYNTSILVHYPEGLSL 819
Qy 814 RKYSTLQORSORSRLACESASSTEVSGALKSTCSINHPIFPENSEVTNITFDV---870
Db 820 SKFDAIKPSRTR-----SSCGDRDSGATNRTTCSINLPVYRSGLTTQFLGTFRTKW 871
Qy 871 DSKASLGNKULLKANVTSENNMPTNKTETQLELPVKYAVTMVVTSHGV-STKYLNFTAS 929
Db 872 DYDWSDRMETITANSNDNGNM---SDMSVRRSIPQFAVELAISLVAEDSVTYLNFSL 928
Qy 930 ENTSRVWQHOYVSNLQORSILSLVLPVRLNQVIMDRPQVTFSENLSSTCHTKERL 989
Db 929 DRGPKPLIYKVNVNGLKLPVSVTLSPCQ-----THVTLTTPHNFMSQ 974
Qy 990 PSHSDFLAEALRKAPVNC-----SIAVCQRIQCDIPFFGQEE 1027
Db 975 EVHHSFISSVHQ--IICLLNKLHFFSPPELSAVQTRITGSLWC-----VSSISTGEI 1026
Qy 1028 FNATLKGNL-----SFDWYIKTSHNHLIVSTAELFNDVSNFTLLPG--OGA 1072
Db 1027 FRSSV--NLMAEAVLQNVKEYESKSYF-BFRDRHVENIS-AELNFTNRYNQSTGLKYN 1082
Qy 1073 FVRSQTEKVEFPVNPPLIVGVSSVGGLLLLALITAAALYKLGFFPKQYKD 1124
Db 1083 PHRSQTEKVEFVIPPBRMLIVCTGVAGVGGFFFLIILILLKCGFFKRNRPD 1134

RESULT 11

Q96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1; AC
DT 01-DSC-2001 (T-EMBLrel. 19, Created)
DT 01-DSC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC008777; AA08777.1; -.
DR InterPro; IPR00413; Integrin_alpha.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWA; 1.
DR Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; P6FF2546E8C632F9 CRC64;
Query Match 23.0%; Score 1350.5; DB 4; Length 1086;
Best Local Similarity 32.2%; Pred. No. 5.6e-30;
Matches 376; Conservative 186; Mismatches 432; Indels 175; Gaps 37;
Qy 1 FNLDENAMTFQ--ENARFGQSVVQLQGSVVRVVGAPQEIIVAAQNGSLYQCYSYTGSC 58
Db 26 YNLDVARGARFSPPRAGRHFGRVQLV-GNGVIVGAPGE---GNSTGSLYQCQSGTGCHL 81
Qy 59 PIRLOVPEAVNMSLGLSLAATTSPPQLLACGPTVHTCSENTYVVKGLCFLFGSNLRQP 118
Db 82 PVTLR-GSNVTSKYLQMTLA--TDP-----103
Qy 119 QKPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMOYSE 178
Db 104 -----TDGS-----ILFAAVQFST 117
Qy 179 EPRHITFEFQNNPNRSLVKPITQLGRTHATGVRKVIKRELLNITNGARKNAFKILI 238
Db 118 SYKTEPFDSDYVVRKDPDALLKHVKHMLLTNTFGAINVYVATEVFRFELGARPDATKVL 177
Qy 239 VIITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFSEKSRQELNFIASKPRDHVQ 298
Db 178 IITDGE--ATDSNIDAAKD-----IIRYIIGIKGHFQTKESQETLHKFASKPASEFVKI 230
Qy 299 VNNFEALKTIONLREKIFATEGTGTGSSSSFEHMSQEGFSAITSNGPLLSLVGSDYD 358
Db 231 LDTFEKLKDLFTLEQKIVYIEGTSKQDLTSFNMELSSSGISADLSRGHAVGAVGAKDW 290
Qy 359 AGGVF-LYTSKEKSTFINMTRVDSMDNDAYLVGAIAA--IILNRNVQSLVLGAPRYOHTGLV 416
Db 291 AGGFLLDKADLDQDITFIGNPLETPVRAGLYGVTVLWPSRQKTSLLASCAPRYQHMGRV 350
Qy 417 AMFR--QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQ 474
Db 351 LLFQEPQGGGHWVQVQTIHGTQIGSYFGGELCGVDVDQDGETELLIGAPLFYGEQRGR 410
Qy 475 VSVCLPLPRGORARWQCDAV--LVGEGOPMGRFGAALTVLGDVNGDKLTDVAIGAPCEED 532
Db 411 VFYI-----QRRQLGFEESVSELOQDGYPLGRFGEAITALTDINGDLVDVAVGAPLEE 464
Qy 533 NRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVLTGVGAQG 592
Db 465 -QGAIVIFNGRHG-GLSPQSPQRIEGTVLSIQWFGRSIHGVKDLGDLGDLADVAVGAES 522

[illegible]

RESULT	ID	Q8HZVO	PRELIMINARY;	PRT;	927 AA.
DB	AC	Q8HZVO;			
DT	01-MAR-2003	(T-EMBLrel. 23, Created)			
DT	01-MAR-2003	(T-EMBLrel. 23, Last sequence update)			
DE	01-MAR-2003	(T-EMBLrel. 23, Last annotation update)			
DI	Lymphocyte function-associated antigen 1 (Fragment).				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;				
OX	Bovidae; Bovinae; Bos.				
NCBI_TaxID=9913;					
[1]					
RN	SEQUENCE FROM N.A.				
RA	Thumrikat P., Kannan M.S., Maheswaran S.K.;				
RT	"Sequence of the alpha subunit of bovine lymphocyte function-				
RT	associated antigen 1.";				
RL	Submitted (OCT-2001) to the				
DR	EMBL; AF440778; AAN63636.1; -.				
FT	NON TER				
NON_TER	1				
NON_TER	927				

Qy	234	FKLLIVITDGEKPGDPLGYEDVPIPEADREGVYRVVLGVGDAPSEKSRQELNTIAASKPPR	293
Db	64	TKVLIITDGG---PPTNTLMRPKTSS-----RSLIGIKNFKTKSQEALHQFASKPVE	116
Qy	294	DHVFQVNNPEALKTIQNLREKIFAJEGTQTCSSSFEHEMSQEGSAAITSNGLLSTV	353
Db	117	EFVKIILDTPEKLDLELQKLIYVLEGTSKQDLTSFNMELESSG:ISADLSEGHGVVGAV	176
Qy	354	GSYDWMAGGVF-LVTSKEKSTFFINMTRVDSMDNDAYLGYAAA-IILNRVOSVLVGAPRYQ	411
Db	177	GAKDWAGGFLDLKADLKSSFTVUGNEPLTVBESRAGYLGYTVTRLPSTRGTMSELLATGAPKYQ	236
Qy	412	HIGLVAMFQ--NTGWNESNANVKTQIGAYFCASLCSVDVDSNGSTDVLVIGAPHYEQ	469
Db	237	HVGRVLLFOQPKRGKGPWSQIQEIDG:QIGSYFGGELCGVDVDRXDETLELLIAAPLYGE	296
Qy	470	TRGQVSVCPPLRGQARWOCDAVLGEQOQPMGRFGAALTVLGDVNGDKLTDVAIGAPG	529
Db	297	QRGRVFYI---QKIQLEFQWVSELOGETGYPLGRFGAA:TAALTDINGDELTDAVGAPL	353
Qy	530	EEDNRGAVLYFHTGSSGIGSPHSQRTAGSKLSPRIQYFGQSLSGGODLTWMDGLVDLTVG	589
Db	354	EE--QGAVYIFNGQG-GLSPRPSQRTEGTMFSGIOWFGRSIHGVKDLGGDGLADVAVG	410
Qy	590	AQGHVLLRSQPVLRVKAINEFNPREVARNVFECNDQVVKGEAG-EVRVCLHVQK--ST	646
Db	411	AEQGVILSRPVPVDIITSVSFPAELPVNHEVECSYSTSNQKEGVNLTVCFOVKSII	470
Qy	647	RDLREGIOISVVTYDLALDSGRPHSRAVNETKNSTRROTQVLGLTQTCETLLKQLPNC	706
Db	471	-----FQGHIVANITYTLQDCHRTSRGLFPGGKHKLIIGNTAVTPV-KSCFVFWFHPIC	525
Qy	707	IEDPVSPIVLRUNFSL---VGTPLS---AFGNLRPVLAEADAQR:LTALPFPKKNCGNDNIC	761
Db	526	IQDLISPINVLSYLSWEEBEGTPDRPALDRDIPILKPSPHLETKEIPEKKNCGDKVC	585
Qy	762	QDLSITSPMSLDCLVGGPREPNVTVVRNDEGDSYRTQVTFPFLDLDSYRKVSTLQON	821
Db	586	EADLKAFSDMRSKILRTPSASLSVALTLRTAEADYAVVQVTLSPFQGLSFRKVEIL--	643
Qy	822	QRSQRWRLACESASSSTEVSALKSTSCSINHPIFPENSEVTNFI:TPDVDSKASLGNKLL	881
Db	644	-KPHSHVPVQCEELPEAAVHS-RALS:CNVSSPIFGEDSWVDIOVMFNTLQKSGWGFIE	701
Qy	882	LKANVTS-----ENNMPRTKNTEPQLELPVKYAVYVMTVSHGVSTKYLNFTASENTS	933
Db	702	LQANVSCNNEDSLLIEDNSATTS-----IPVMYPINVLTKQDENSTLYISFTPKSPRI	754
Qy	934	RVMOHQYQV-----SNLCQSRSLPTSLVFLVPVRLNQTVI---WD---RPQVTFSS-ENLSST	982
Db	755	HVVKHIYQVRIQPSNYONMP-PLBALVRVPRVHSEGLITHKWSIQMBPPVNCSPRNLESP	813
Qy	983	CHTKERLPSHSDFLAELRKAPVNCSTAVCQRIODDIPFFGIGEEFNATLKNLSFDWYI	1047
Db	814	SDEAE-----SCSRGT--EFRCPIDF---ROEILVQVNGMVELRGTI	850
Qy	1043	KTSHNHLIIVSTABILFENDSVFTLLPQOGAFVRSQ:TEKKBPEFVNPPLPIVSGSSVGLL	1102
Db	851	KAS-SMLSLSSLSAISNSSKHPHLHGRNASM-AQVVMKVVDLVYEKEMLYLVLSGIGGL	908
Qy	1103	LLLALITAALYKLGFFKR	1120
Db	909	LLLEFIALYKVGFFKR	926

RESULT 13					
O88340					
ID	O88340	PRELIMINARY:	PRT:	1167 AA.	
AC	O88340;				
DT	01-NOV-1998	(TREMBLrel. 08,	Created)		
DT	01-NOV-1998	(TREMBLrel. 08,	Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23,	Last annotation update)		

```
DE Integrin alpha E1. (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079 (1997).
DR EMBL; AF020045; AAC23662.1;
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
DR PROSITE; PS02334; VWF; 1.
SQ SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match 19.8%; Score 1163.5; DB 11; Length 1167;
Best Local Similarity 29.1%; Pred. No. 3.6e-76;
Matches 359; Conservative 211; Mismatches 458; Indels 205; Gaps 43;

Qy 1 FNLDTENA--MTFOENARGQSVVQLQGRVVVVGAPQEIIVAAHQGS-----LYQCDY 52
Db 1 FNLDTENA--MTFOENARGQSVVQLQGRVVVVGAPQEIIVAAHQGS-----LYQCDY 52
Qy 20 FNMVDVWAVTALQPGAPVLSLLHLDPSN-----NQTCLLVARRSSNNTAALYRCAL 74
Db 20 FNMVDVWAVTALQPGAPVLSLLHLDPSN-----NQTCLLVARRSSNNTAALYRCAL 74
Qy 53 STGCEPIRLQVPVEAVNMSLGLSLAAT--SPQLLAC-GPTVHQTCSNTYVYKGLCFL 109
Db 53 STGCEPIRLQVPVEAVNMSLGLSLAAT--SPQLLAC-GPTVHQTCSNTYVYKGLCFL 109
Qy 75 SI-SPDEIACQ-PVEHICMPKRGYQVTLVGNHNGVLVCIQVQARKPSLSNLSLTGACSL 132
Db 75 SI-SPDEIACQ-PVEHICMPKRGYQVTLVGNHNGVLVCIQVQARKPSLSNLSLTGACSL 132
Qy 110 FGSNLROQPKFPALRG-----C-----PQE 131
Db 110 FGSNLROQPKFPALRG-----C-----PQE 131
Qy 133 LTPNLDLQAAQYFSDLEGLDPAHVDGSGDYCRSGSGTGEKKSARRRTVEEBDEED 192
Db 133 LTPNLDLQAAQYFSDLEGLDPAHVDGSGDYCRSGSGTGEKKSARRRTVEEBDEED 192
Qy 132 DSDIAFLIDGSGSIIPDFRRMKFVSTVMEQL--KSKTLFSLMQYSEEFRIHFTPEKF 189
Db 132 DSDIAFLIDGSGSIIPDFRRMKFVSTVMEQL--KSKTLFSLMQYSEEFRIHFTPEKF 189
Qy 193 GTEAIVLDGSGSIEPDSDFQAKNFISTMRNFYKCFECNFAVQYGAIVQTFEFLQES 252
Db 193 GTEAIVLDGSGSIEPDSDFQAKNFISTMRNFYKCFECNFAVQYGAIVQTFEFLQES 252
Qy 190 QNNPNRSLVKPITQLLGRTHATGVRKVIRELNIITNGARKNAFKLIVITDGEKFGDP 249
Db 190 QNNPNRSLVKPITQLLGRTHATGVRKVIRELNIITNGARKNAFKLIVITDGEKFGDP 249
Qy 253 RDINASLAKYQSIQVQKEVTKTASAMQHVLDNIFIPSRGSRKALKVMVLTGDIQFQDP 312
Db 253 RDINASLAKYQSIQVQKEVTKTASAMQHVLDNIFIPSRGSRKALKVMVLTGDIQFQDP 312
Qy 250 LGVEDVTPADREGVIRYVGVGDAFRSEKSRQBLNTIASKPPRDHVFQVNNFEALKTIO 309
Db 250 LGVEDVTPADREGVIRYVGVGDAFRSEKSRQBLNTIASKPPRDHVFQVNNFEALKTIO 309
Qy 313 LNLTTVINSKPMQGVVRFALGDAFKNNNTYRELKLIASDPKEAHTFKVTNYSALDGLL 372
Db 313 LNLTTVINSKPMQGVVRFALGDAFKNNNTYRELKLIASDPKEAHTFKVTNYSALDGLL 372
Qy 310 NQLREKFAIEGTGTGSSSEFHEHMSQEGFSAATTSNGP--LLSTVSGYVDWAGGVFLY-TS 367
Db 310 NQLREKFAIEGTGTGSSSEFHEHMSQEGFSAATTSNGP--LLSTVSGYVDWAGGVFLY-TS 367
Qy 373 SKLQORIVHHEGT---VGDALQYLAQTGFSAQILDKGVLLGTGAFNWSGGALLYSTQ 429
Db 373 SKLQORIVHHEGT---VGDALQYLAQTGFSAQILDKGVLLGTGAFNWSGGALLYSTQ 429
Qy 368 KEKSTFTINMT-RVDS-DMNDAYLGYAAAIIILNRNVQSLVLGAPRYOHIGLVAFRONTGM 425
Db 368 KEKSTFTINMT-RVDS-DMNDAYLGYAAAIIILNRNVQSLVLGAPRYOHIGLVAFRONTGM 425
Qy 430 NGRCCFLNQAKEDSRVQVSYLGYSLAVLHKAHGISYVAGAPRKLGRGAVFELRKEDR- 488
Db 430 NGRCCFLNQAKEDSRVQVSYLGYSLAVLHKAHGISYVAGAPRKLGRGAVFELRKEDR- 488
Qy 426 WESNA----NVKGTQIGAFYFASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQVSCVPLR 482
Db 426 WESNA----NVKGTQIGAFYFASLCSVDVDSNGSTDLVLIGAPHYYEQTGRGGQVSCVPLR 482
Qy 489 -EEDAFVRRIEGEQMGVYFSGVLCFVDDIMDGTDFLLVAAPFVHIRGEEGRVYVYQVE 547
Db 489 -EEDAFVRRIEGEQMGVYFSGVLCFVDDIMDGTDFLLVAAPFVHIRGEEGRVYVYQVE 547
Qy 483 GQARWQCDVLYEQCGQPCRCFCAALTVLGDVNGDKLTDAICAP-----GEEDNRGA 536
Db 483 GQARWQCDVLYEQCGQPCRCFCAALTVLGDVNGDKLTDAICAP-----GEEDNRGA 536
Qy 548 -QDASFSLAHTLSHPGLTNSRFGFAVAVGDINDQKFTDVAIGAPLEGFGAGDCASYGS 606
Db 548 -QDASFSLAHTLSHPGLTNSRFGFAVAVGDINDQKFTDVAIGAPLEGFGAGDCASYGS 606
Qy 537 VYLFHGTSGSGISPSHSQRAGSKLSPRLQYFGQSLSGGGDLTMDGLVLDLTGVAQGHVLL 596
Db 537 VYLFHGTSGSGISPSHSQRAGSKLSPRLQYFGQSLSGGGDLTMDGLVLDLTGVAQGHVLL 596
```

```
Db 607 VYIYNHSG-GLYDSSQQIRASSVSGGLDFNGDGLADITVGSRDVAVV 665
Qy 597 LRSQPVLRKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
Db 666 LRSRPVVDLTVSMTFTP-----DALPMVFIGKM--DNLCCFEVDSSVASEPGLREM 715
Qy 654 QIQSVVYDLDALDSGRPHSRVFNETKNSRQTOVLGLTQTC----- 696
Db 716 FLNFTVDVAV-----TKQRQLQCEDESSGQCLRWKNGSFLCEHFMLI 760
Qy 697 ETLLQLPNCIEDPVPISVLNLSLVGTPLSAFGNLR---PVLAEADQRLFTALF--P 750
Db 761 STEEL-----CEDCFNSITIKVTYE-----FOTSGGRRDYPNPTL--DHYKEPSAIFQLP 809
Qy 751 FEKNCGNNDNICODLSITFSFMSLCLVVGGRPFNVTVTVNRDGEDSYRTVTFPPFLD 810
Db 810 YEKDCNKNVFCIAEIQLTN--ISQQLVVGTVKEVTMNIISLTNSGEDSYTNMALNYPN 868
Qy 811 LSVKYSTLQONORSQSWRLACASSTEVSGAKSTSCSINHPIPPENSEVTNITFDV 870
Db 869 LQPKKI-----QKVPSPDVQCDPKPV--ASVLVNMCKIGHPII-KRSSVNVSVTQWL 918
Qy 871 DSKASLGNKLLKCANVTSNNMPRTNKTETFOLELPVKYAVYVTVTSHGVSTKYLNTASE 930
Db 919 EESVFNRTADITVTIISNKSLEARETR---SLOFRHAFIAVLRS--PSVMYMN--TSQ 971
Qy 931 NTSRVNQHVQVSNLQGRSLPISLVLFPVRLNQTIVMDRPOVTPSPENLSST-----CHT 985
Db 972 SPDSHKKEFFNVHGENLFGAVFQLOICVPIKLODP-----QIVRVKNLTQDHTECTQ 1025
Qy 986 KERLPSPHFLAELRAKAPVNVNCSIAVCQRICQDIPFGIOSEFNATLKGNSLFDWYIKTS 1045
Db 1026 SQBPACGSDPVQVQKWHWSVCAI-----TSNKENVTVAEISVG 1065
Qy 1046 HNHLLIVSTA-----EILFNDSVFTLLPGQAFVRSQTETKVEFP-----EVNPLPLIV 1095
Db 1066 HTKQLLRDSELPIGEISFNKSLYEGLNAB-----NHRTKITVIFLKEETRSPLII 1119
Qy 1096 GSSVGGLLLALITALYALYKLGFFKROYKDMNSE 1128
Db 1120 GSSIGGLLVVLIILFKCGFFKRYQQLNLE 1152

RESULT 14
O88341 PRELIMINARY; PRT; 1167 AA.
AC O88341;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079 (1997).
DR EMBL; AF020046; AAC23663.1;
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 4.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
```

```
DR SMART; SM00191; Int_alpha; 4.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.  
DR PROSITE; PS50234; VWFA; 1.  
FT NON_TER  
SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBEC CRC64;  
  
Query Match  
Best Local Similarity 27.8%; Pred. No. 6.7e-73;  
Matches 335; Conservative 204; Mismatches 423; Indels 244; Gaps 37;  
  
QY 47 LYQDYSTGSCSEPIRLQVPVEAVNMSLGLSLAAATSPOLLACGPTVH----- 94  
DQ |||||  
Db 67 LYQC-----ATSIIDPICCPVEHILMPKGRYQGV 98  
QY 95 -----QTCSENVYKGLFGLGSLNRQKQKPEALRG----- 127  
Db 99 LVNRHNGVLVCIQOSRKPRNSLSETGACSLTLPNLOLQAQYFSDLEGVLDGASVNS 158  
QY 128 -----CPQED-----SDIAFLIDGSGIIPHDFRMKFV 157  
Db 159 GDYSGKSGTGEETKSARLRQRAVEEEDBEAEAGTEIAVLGSGSIEPSDFQAKDFI 218  
QY 158 STWMBOQL--KSKTLFSLMQVSEEPRIHFTFEFQNNPNRSLVKPIITQLLGRTHATGV 215  
DQ |||||  
Db 219 STMRNFYKCEPCNFALVOYGVITQTEFDLLDSRDINASLAKVQSIQVQKVEVTKASAM 278  
QY 216 RVIRELLNITNGARKNAKILVITDCEKGDPLGYEDVPEADREGVIRVIGVGDAF 275  
DQ |||||  
Db 279 QHVLONIFIPSGSRKKALKVMVLTGDIFRDLPLNLTTVSSSKMQGVVPAIGVGNAP 338  
QY 276 RSEKSRQELNTASKPPDRHVRQVNNFEALKTIONLREKIPAEIGTQTGSSSSPEHMS 335  
DQ |||||  
Db 339 ENNNYRELKLASDPKAAHTKVTNYSALDGLSKLQRIIHMEGT---VGDTLQYQLA 395  
QY 336 QGFGSAITSNGP--LLSTVSGVDWAGGVPLY--TSKEKSTFINMTRVDSMDND---YLGY 390  
DQ |||||  
Db 396 QTGFSAQILDKQVLLGTGAFNWSGALLYNTQNGRFLNQT--AKEDFRAAQSYLGY 454  
QY 391 AAAILLRNRVQSLVIGAPYQHIGLVAMPFRONTGWESNANVKGTOIGAYGASCSVDV 450  
DQ |||||  
Db 455 SVAALHKAGVSVAGAPRHLRGAVFELQKEDGETFMRRTEGEQMGSYFGSLVCPVDI 514  
QY 451 DSGNSTDLVLIGAPHYETROGQVSVCPPLRGQARWQCDVLYGEOQOPWGRFGAALT 510  
DQ |||||  
Db 515 NWDGIDTDLVLAAPPHIRGEGRVYVYRV--HEQDAPFLSVLTLSGYPGLTSSRGFAMA 573  
QY 511 VLGDVNGDKLTDVAIGAP-----GBEDNRGAVYLFHGTSGSIGSPHSQRSIAGSKLGR 564  
DQ |||||  
Db 574 AVGDINQDKFTDVAIGAPLEGFGAGDGASGVYIYNGHSG--GLHASPSQIIRASSVALG 632  
QY 565 LQYFGOSLGGQDLWDGLVLTGAGQHVLLRQPVLRVKAIMEFNPREVARNVPECN 624  
DQ |||||  
Db 633 LYFGMSVSGGLDFSGDDLDATVGSQDVAVVLRSPVVDLTVSMFTTDPALP----- 685  
QY 625 DVVYKKEAGEVRVCLHVQKS---TRDLREGIOISVTVTYDIALDGRPHSRVAFNETGN 681  
DQ |||||  
Db 686 ----MAFKDMVELCFKVDSAVPSEPLGRGSLNFTVDV-----TKQ 727  
QY 682 STRRQ-----TQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNFS 721  
DQ |||||  
Db 728 KQRLQCADRSQCSCSLMKWSGGSSLCHEFLISTEEL-----CEDDCFSNITIKVSYE 781  
QY 722 LVGTPLSAGNLRPVLAEDAQLRFTALF--PEKNGCNDNI CODDLSITFSMSLDCLVY 779  
DQ |||||  
Db 782 F-QTSEERNHPNPIL--DHYKEPSAIFQLPYEKDCKNVFCFAETQLTALISQOD-LVV 837  
QY 780 GGPREFNVTVYRNDGEDSVYRQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTE 839  
DQ |||||  
Db 838 GITKEVTMISLNSGSDSYMNTMALNYPNNQFKKI-----QKPLSPDIQCDPKPV- 890  
QY 840 VRGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKT 899  
DQ |||||
```

```
QY 108 FLFGSNLROQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQLK-- 165
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
138 SNVSTFETVKAAP--SVQEC-KTQLDIVILDGSGNSIYP--WESVTAFLNLRNMDIG 193
QY 166 KSKTFLSLMOYSEEFRIHTFKEFQNNPNRSLVUKBITQLLG-RHTATGVKRVIRELN 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
194 POOTQVGIYQYQGVVHFEYLNSTYTEEVMDAALRIQRGGTQMTALGIDITAREEFT 253
QY 225 ITNGARKNAFKILIVITDGEKFDGLGEDVPEADREGVIRYVIGVDAP-----RSEK 279
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
254 EAHGARGVQKVMVIVTDGESH-DNYRLQEVTDKCEDENIQRPALIALGYSRGNLSTEK 312
QY 280 SRQELNTIASKPRDRHVFQVNFPEALKTIQNOLREKIFAIEGTQTGSSSFHEMSQEGF 339
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
313 FVEEIKSIAKPKTEKHFNVSDALVTVIVEALGERIFALEATTDOQAAAFEMEMSQAGF 372
QY 340 SAAITSNGLPLSTVGSYDHWAGGVFLYTSKEKSTFINMTRVDSMDND-----AYLGYAAAI 394
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
373 SAHYSQDWYMLGAVGAYDWNGTVMYVKDSIDIPSNDTFRDRHSEKIEPLAAYLGYTVNS 432
QY 395 ILNRVQSLVLGAPRYOHLGIVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNG 454
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
433 ALTPGGVLYIAGQPRVNHHTGQVVIYKMEGREVKVQLRLGEBQIGSYFGGVITTTIDINRDS 492
QY 455 STDVLVLIGAPHYY--EQTRGGQVSVCLPRG-----QRARWQCDVLYGE-----QG 499
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
493 FTDLLLVGAPTYMGTEKESQGVYVYALNKTPEYQMSLEPIKQTCSPKXHDTCVKLN 552
QY 500 QPWG-RFGAALTVLGVNNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAG 558
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
553 EPCGARFGTAIAAVKDLNLDGVNDIVIGSPLEDDHRGAVVIYHG-HGNTISKYTORIAS 611
QY 559 SKLSPRLOYFGOSLGGQDLTMDGLDVLTVGAQGHVLLRSQPLVLRKAIEMEFNPREVAR 618
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
612 GGDGEKVFPGSGVHGMNDLNDGLDVTIGLGGAAFWSDVAENVVSMQFMPKSI-- 669
QY 619 NVFECNDQVVKGEAGEVRVCLH--VQKSTRDLREGQIQSVVTVYDIALDSGRPHSRAPV 676
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
670 NIOQNCQINKRK-----TICINATICFKTRUKSKEDIFESSLOYWITLDAQOISRSLF 724
QY 677 NETKNSRRTQTVLGTQTCETLK-----LQPNCTIEDPVSPVLRNLSVLTGTPLSAFG 731
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
725 TETH---ERKMQNITIKGSECIKHNFMLDKPD-PQDSVK-VLLEFNFS---DPESG-- 774
QY 732 NLRPLAEDAQRLLFTALFPKNCNGNDICODDLSTFSF-----MSLDCLVUGGPREN 786
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
775 ---PVLDSNLPNSISEYIPTKDCGAKNKCISDLALNVKASIAGDSSSPFIVKSRNDRFT 831
QY 787 VTVVRNDGEDSVRTQVTFPFLDLSYRKVSTLQONQORSQSWRLACESASSTEVSGALKS 846
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
832 IQLSIKNKKDSAYNTRALVQYSPNIIIFAGIEDIOXD-----SCESN-----HN 874
QY 847 TSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENMP-----RTNKTFFOL 902
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
875 ITCKVGYPLKPEEISFKISQFNASYLLEN-ATVDVYVATSDSEBPPETLRDNRG--QV 931
QY 903 ELPVKYAVVMVYTS-----HGVS TKYLNFTASENTSRVMHQ-----YQVSNLQORSLP-I 952
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
932 TIPVKYEVGLIFVSVFKEHHVIAANETIPTAINTTEQIGDEVTLHYRIKGEHFPFMPNL 991
QY 953 SLVFLVP--VRLNQTVMORPOVTFSENLSSTCHTKE-----RLPSHSDFLAELRKAP-- 1003
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
992 TIQILYPDVTTAKNTLLY--ITTLHSQNAICKSSYPVDHLKIGSGKSYVLPKIKEPTK 1048
QY 1004 --VVNCISIAVCORIQCDIPFGIQEEFNATLKNLSFDW---YIKTSHNHLIIVSTAAIL 1058
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1049 DTIMECDTFCASINCALAPSDI-SQVNVSLR-----VWKPTIIKASIHSLTIVVKALLR 1102
QY 1059 FNDSVFTLLPGQGAFVRSQETKTEKPFEPVNPPLP--IVGSSVGGLLLLALITAILYKLG 1116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1103 SENSSLIL---RNDHOKLETMIKISKEPPPGSVPLWVPLPSIFAGLILALLIFALWKAG 1159
```

QY 1117 FFKROYKDMSE 1128
Db 1160 FFKRPLKKMEK 1171

Search completed: November 25, 2003, 14:20:40
Job time : 38.1068 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:08:29 ; Search time 36.2978 Seconds
(without alignments)
4971.979 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19JUN03.*

1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5868	99.9	1153	19	AAW65090 Human Beta-integrin
2	5868	99.9	1153	21	AAW65090 Human CD11b protei
3	5868	99.9	1153	23	ABG61469 Human Beta2 integr
4	5868	99.9	1153	23	AAU80252 Human integrin 1 a
5	5868	99.9	1153	23	AAO34428 Integrin Mac-1 alp
6	5858	99.7	1153	11	AAO4136 Alpha subunit of M
7	3464	59.0	1163	11	AAW07120 p150.95 alpha subu
8	3450	58.7	1163	19	AAW65091 Human Beta-integrin
9	3450	58.7	1163	21	AAW07361 Human CD11c protei

10	3450	58.7	1163	23	ABG61470 Human Beta2 integr
11	3448	58.7	1163	24	ABU07406 Protein differenti
12	3417	58.2	1161	16	AAW78166 Human beta-2 integ
13	3417	58.2	1161	18	AAW23049 Human beta 2 integ
14	3417	58.2	1161	19	AAW72825 Human alpha-d. Ho
15	3417	58.2	1161	19	AAW65089 Human Beta-integrin
16	3417	58.2	1161	19	AAW57491 Human Beta2 integr
17	3417	58.2	1161	20	AAW73342 Human alphad prote
18	3417	58.2	1161	21	AAW07359 Human alpha d clon
19	3417	58.2	1161	23	ABG61468 Human Beta2 integr
20	3401.5	57.9	1161	18	AAW23064 Human beta 2 integ
21	3401.5	57.9	1161	19	AAW72837 Human alpha-d deri
22	3401.5	57.9	1161	19	AAW65106 Human Beta-integrin
23	3401.5	57.9	1161	20	AAW73343 Human alphad prote
24	3401.5	57.9	1161	21	AAW07376 Human alpha d prot
25	3401.5	57.9	1161	23	ABG61485 Human Beta2 integr
26	3241.5	55.2	1161	16	AAW78169 Rat alpha-d subuni
27	3239.5	55.1	1161	18	AAW23082 Rat beta 2 integrin
28	3239.5	55.1	1161	19	AAW72834 Rat alpha-d #1. R
29	3239.5	55.1	1161	19	AAW60004 Rat alpha d polype
30	3239.5	55.1	1161	21	AAW07374 Rat alpha d protei
31	3239.5	55.1	1161	23	ABG61483 Rat Beta2-integrin
32	3232.5	55.0	1161	19	AAW65104 Rat beta-integrin
33	3232.5	55.0	1161	20	AAW73345 Rat alphad protein
34	3228	54.9	1161	18	AAW23061 Mouse beta 2 integ
35	3228	54.9	1161	19	AAW72836 Mouse alpha-d #2.
36	3228	54.9	1161	19	AAW65103 Mouse beta-integri
37	3228	54.9	1161	19	AAW60003 Mouse alpha d poly
38	3228	54.9	1161	20	AAW73347 Mouse alphad prote
39	3228	54.9	1161	21	AAW07373 Mouse alpha d prot
40	3228	54.9	1161	23	ABG61482 Mouse Beta2 integr
41	3224	54.8	1161	16	AAW78168 Mouse alpha-d subu
42	3218.5	54.8	1151	18	AAW23059 Rat beta 2 integrin
43	3218.5	54.8	1151	19	AAW72834 Rat alpha-d #2. R
44	3218.5	54.8	1151	19	AAW65101 Rat beta-integrin
45	3218.5	54.8	1151	19	AAW60001 Rat alpha d polype

ALIGNMENTS

RESULT 1

AAW65090

ID AAW65090 standard; Protein; 1153 AA.

AC AAW65090;

DT 28-SEP-1998 (first entry)

DB Human Beta-integrin CD11b subunit protein.

KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
KW rheumatoid arthritis.

XX Homo sapiens.

XX OS

FN US728533-A.

XX

PD 17-MAR-1998.

XX

PF 07-JUN-1995; 95US-0485618.

XX

PR 07-JUN-1995; 95US-0485618.

PR 23-DEC-1993; 93US-0173497.

PR 05-AUG-1994; 94US-0286889.

PR 21-DEC-1994; 94US-0362652.

XX (ICOS-) ICOS CORP.

PA

XX Gallatin WM, Van DER VIEREN M;

PI

XX

DR	WPI; 1998-206565/18.	
XX	Screening assay for modulators of integrin binding - using	
PT	immobilised or labelled alpha-d polypeptide, useful for, e.g.	
PT	treating type-1 diabetes	
XX		
PS	Example 5; Fig 1A-D; 106pp; English.	
XX		
CC	This sequence represents a human beta-integrin CD11b subunit which is	
CC	used to describe a method for identifying compounds that modulate the	
CC	interaction of the beta-integrin alpha-d subunit with a binding partner	
CC	of alpha-d which involves contacting an alpha-d polypeptide with an	
CC	alpha-d binding partner, one of which is immobilised and the other of	
CC	which is labelled, in the presence of a test compound, and determining if	
CC	the compound affects binding between the alpha-d polypeptide and alpha-d	
CC	binding partner, where the alpha-d polypeptide is alpha-d or its fragment	
CC	comprising the cytoplasmic, transmembrane or extracellular domain of	
CC	alpha-d. Compounds that modulate alpha-d binding could be used to treat	
CC	diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,	
CC	asthma, psoriasis, lung inflammation, acute respiratory distress	
CC	syndrome and rheumatoid arthritis.	
XX		
SQ	Sequence 1153 AA;	
	Query Match 99.9%; Score 5868; DB 19; Length 1153;	
	Best Local Similarity 99.8%; Pred. No. 0;	
	Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
QY	1 FNLDTENAMTFQENARGFGQSVVLOGSRVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 60	
DB	17 FNLDTENAMTFQENARGFGQSVVLOGSRVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 76	
QY	61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 120	
DB	77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 136	
QY	121 FPALRGCPQEDSDIAFLDGSIIIPHDFRRMKFVSTVMEOLKSKTLFSLMQLYSEEF 180	
DB	137 FPALRGCPQEDSDIAFLDGSIIIPHDFRRMKFVSTVMEOLKSKTLFSLMQLYSEEF 196	
QY	181 RIHFTKFEFONNPNRSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 240	
DB	197 RIHFTKFEFONNPNRSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 256	
QY	241 TDGEKFGDPLGYEDVPEADREGVIRVVGADAFRSEKSRQELNTIASKPRDHVQVN 300	
DB	257 TDGEKFGDPLGYEDVPEADREGVIRVVGADAFRSEKSRQELNTIASKPRDHVQVN 316	
QY	301 NFALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLSVTGSDWAG 360	
DB	317 NFALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLSVTGSDWAG 376	
QY	361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420	
DB	377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYOHIGLVAMFR 436	
QY	421 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTREGQVSCPL 480	
DB	437 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTREGQVSCPL 496	
QY	481 PRGORARWQCDVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540	
DB	497 PRGORARWQCDVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556	
QY	541 HGTSGSGISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600	
DB	557 HGTSGSGISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616	
QY	601 PVLKRVKAIMBNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGIQSVVT 660	
DB	617 PVLKRVKAIMBNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGIQSVVT 676	
QY	661 YDLALDSGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRNF 720	

DB	677 YDLALDSGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRNF 736	
QY	721 SLVGTPLSAFQNLPRVLAEDAQRFLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780	
DB	737 SLVGTPLSAFQNLPRVLAEDAQRFLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796	
QY	781 GPREFNVTVVRNDEGSYRTQVTFPPFLDLSYRKVSTLQNRORSRWRLACESASSTEV 840	
DB	797 GPREFNVTVVRNDEGSYRTQVTFPPFLDLSYRKVSTLQNRORSRWRLACESASSTEV 856	
QY	841 SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900	
DB	857 SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916	
QY	901 QLELPVKVAVMVVTSHGVSSTKYLNTASNTSRVMQHYQVSNLQSRSLPISLVFLVPV 960	
DB	917 QLELPVKVAVMVVTSHGVSSTKYLNTASNTSRVMQHYQVSNLQSRSLPISLVFLVPV 976	
QY	961 RLNQTVIWRDPQVTFSENLSTCHTKERLPSPHSDFLAELRKAPVNVNCSIAVCQRIQCIP 1020	
DB	977 RLNQTVIWRDPQVTFSENLSTCHTKERLPSPHSDFLAELRKAPVNVNCSIAVCQRIQCIP 1036	
QY	1021 FFGIOEEFNATLKGNLSPDWDYIKTSHNHLIVSTAEILLFNDSVFTLLPGQCAFVSOTET 1080	
DB	1037 FFGIOEEFNATLKGNLSPDWDYIKTSHNHLIVSTAEILLFNDSVFTLLPGQCAFVSOTET 1096	
QY	1081 KVEPEVENPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYKQKMMSEGGPPGABPQ 1137	
DB	1097 KVEPEVENPPLIIVGSSVGGLLLLALITAAALYKLGFFKRYKQKMMSEGGPPGABPQ 1153	
RESULT 2		
AAB07360		
ID	AAB07360 standard; Protein; 1153 AA.	
XX		
AC	AAB07360;	
XX		
DT	17-JAN-2001 (first entry)	
XX		
DE	Human CD11b protein sequence.	
XX		
KW	Human; macrophage infiltration inhibition; alpha d integrin;	
KW	leukocyte integrin; Leu-CAM; leukointegrin; immune response;	
KW	inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;	
KW	atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;	
KW	lung inflammation; acute respiratory distress syndrome; Crohn's disease;	
KW	rheumatoid arthritis; central nervous system injury; CD11b.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200029446-A1.	
XX		
PD	25-MAY-2000.	
XX		
PF	16-NOV-1999; 99WO-US27139.	
XX		
PR	16-NOV-1998; 98US-0193043.	
PR	08-JUL-1999; 99US-0350259.	
XX		
PA	(ICOS-) ICOS CORP.	
XX		
PI	Gallatin MW, Van Der Vieren M;	
XX		
DR	WPI; 2000-387751/33.	
PT	Use of novel anti-alpha integrin d monoclonal antibodies to inhibit	
PT	macrophage infiltration and reduce inflammation at central nervous	
PT	system injury sites -	
XX		
PS	Example 5; Fig 1; 270pp; English.	
XX		
CC	Integrins are a class of membrane-associated molecules that participate	

in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins. Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha d. The Alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

Query Match 99.9%; Score 5868; DB 21; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDENAMTFQENARFGQSVWOLQGSRVVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDENAMTFQENARFGQSVWOLQGSRVVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVAVNMSLGLSLAATSPOLLACGPTVHTQTSNTYVKGCLFPLFGSNLRQOPQK 120
Db 77 RLQVPVAVNMSLGLSLAATSPOLLACGPTVHTQTSNTYVKGCLFPLFGSNLRQOPQK 136

Qy 121 FPEALRCGPQSDSDIAFLIDGSGIIPHDFRMKEFVSTWMEQKSKTLFSLMOYSEEP 180
Db 137 FPEALRCGPQSDSDIAFLIDGSGIIPHDFRMKEFVSTWMEQKSKTLFSLMOYSEEP 196

Qy 181 RIHFTFEFQNNPRLSLVPIITOLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFEFQNNPRLSLVPIITOLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 256

Qy 241 TDGSKFGDPLGYEDVPEADREGVIRVYGVGDAPRSEKSRQELNTIASKPPRDHVFQVN 300
Db 257 TDGSKFGDPLGYEDVPEADREGVIRVYGVGDAPRSEKSRQELNTIASKPPRDHVFQVN 316

Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSPHEMSQEGFSAIITNGPLISTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSPHEMSQEGFSAIITNGPLISTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINMTVRVSDMDNDAYLGAAAAIILNRVOSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRVSDMDNDAYLGAAAAIILNRVOSLVLGAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQSVCP 480
Db 437 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQSVCP 496

Qy 481 PRGORARWQCDVLYGSGQGPWGRFGAALTVLGVNNGDKLTVDVAIGAPBEDNREGAVYLF 540
Db 497 PRGORARWQCDVLYGSGQGPWGRFGAALTVLGVNNGDKLTVDVAIGAPBEDNREGAVYLF 556

Qy 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTGAGQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTGAGQGHVLLRSQ 616

Qy 601 PVLRVKAIMFENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMFENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736

Qy 721 SLVGTPLSAFNGLRPVLAEADAQRLFTALFPFEKXKCGNDNICQDDLSITTFMSLDCLVVG 780

Db 737 SLVGTPLSAFNGLRPVLAEADAQRLFTALFPFEKXKCGNDNICQDDLSITTFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDGSDSVRTQVTFEPFLDLSYRKVSTLQNRORSQSWRLACSSASSTEV 840

Db 797 GPREFNVTVVRNDGSDSVRTQVTFEPFLDLSYRKVSTLQNRORSQSWRLACSSASSTEV 856

Qy 841 SGALKSTSCSINHPIPIPEENSEVTFNITFDVDSKASLGNKLLKLLKANTVTSENNMPTNTKTEF 900

Db 857 SGALKSTSCSINHPIPIPEENSEVTFNITFDVDSKASLGNKLLKLLKANTVTSENNMPTNTKTEF 916

Qy 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASNTSRVMOHQYQVSNLQORSPLISLVFLVPV 960

Db 917 QLELPVKYAVYVMVTVSHGVSTKYLNFTASNTSRVMOHQYQVSNLQORSPLISLVFLVPV 976

Qy 961 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSPSHDSFLAELRKAPVNVCSIAVCORIOCDIP 1020

Db 977 RLNOTVIMDRPQVTFSENLSSTCHTKERLPSPSHDSFLAELRKAPVNVCSIAVCORIOCDIP 1036

Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLILIVSTAELFNDSVFTLLPGQCAFVRQSOTET 1080

Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLILIVSTAELFNDSVFTLLPGQCAFVRQSOTET 1096

Qy 1081 KVEPFEVPNPPLIVGSSVGGLLLLALITAALYKLGFFKRYKDMHSEGGPPGAEPO 1137

Db 1097 KVEPFEVPNPPLIVGSSVGGLLLLALITAALYKLGFFKRYKDMHSEGGPPGAEPO 1153

RESULT 3
ABG61469
ID ABG61469 standard; Protein; 1153 AA.
XX AC ABG61469;
XX DT 27-AUG-2002 (first entry)
XX DE Human Beta2 integrin alphaCD11b subunit.
XX KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX OS Homo sapiens.
XX PN WO200230980-A2.
XX PD 18-APR-2002.
XX PF 15-OCT-2001; 2001WO-US32059.
XX PR 13-OCT-2000; 2000US-0688307.
XX PA (ICOS-) ICOS CORP.
XX PI Gallatin WM, Van Der Vieren M;
XX DR WPI; 2002-463260/49.
XX PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor
PT recovery, inhibiting locomotor damage, limiting locomotor impairment,
PT or limiting autonomic and sensory dysfunction following spinal cord
PT injury -
XX PS Example 5; Page 191-194; 270pp; English.
XX CC The invention relates to promoting locomotor recovery, inhibiting
CC locomotor damage, limiting locomotor impairment, or limiting autonomic
CC and sensory dysfunction following spinal cord injury by administering an

anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a spinal cord injury victim. The method also involves the use of a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule, vascular cell adhesion molecule). The method is useful for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury. In particular, the spinal cord injury comprises compression of the spinal cord. The antibodies are also useful for reducing inflammation at the site of a central nervous system injury. The specification also details the identification of Beta2 integrin alpha2 cDNAs and proteins, for use in raising the antibodies. Beta2 integrins are implicated in diseases such as LAD (leukocyte adhesion deficiency, inflammatory response, diabetes, multiple sclerosis, arthritis, graft atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, immune complex alveolitis and leukaemia. The present sequence is a Beta2 integrin alpha subunit and leukaemia. The present sequence is a Beta2 integrin alpha subunit sequences.

CC anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a
CC spinal cord injury victim. The method also involves the use of
CC a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion
CC molecule, vascular cell adhesion molecule). The method is useful for
CC promoting locomotor recovery, inhibiting locomotor damage, limiting
CC locomotor impairment, or limiting autonomic and sensory dysfunction
CC following spinal cord injury. In particular, the spinal cord injury
CC comprises compression of the spinal cord. The antibodies are also useful
CC for reducing inflammation at the site of a central nervous system injury.
CC The specification also details the identification of Beta2 integrin
CC alpha2 cDNAs and proteins, for use in raising the antibodies. Beta2
CC integrins are implicated in diseases such as LAD (leukocyte adhesion
CC deficiency, inflammatory response, diabetes, multiple sclerosis,
CC arthritis, graft atherosclerosis, inflammatory bowel disease,
CC Crohn's disease, ulcerative colitis, immune complex alveolitis
CC and leukaemia. The present sequence is a Beta2 integrin alpha subunit
CC sequences. The present sequence is a Beta2 integrin alpha subunit
CC sequences.

XX
SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 23; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCLFGLSNLRQPPQK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVYKGLCLFGLSNLRQPPQK 136
QY 121 PPEALRGCPQEDSDIAFLIDGSGSIIPDFRMRKGFVSTVMEQLKSKTFLSLMOYSEEF 180
DB 137 PPEALRGCPQEDSDIAFLIDGSGSIIPDFRMRKGFVSTVMEQLKSKTFLSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNSLVKPTITQLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNSLVKPTITQLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFRSEKSRQELNTIAKPPRDRHVQVN 300
DB 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFRSEKSRQELNTIAKPPRDRHVQVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTSSESSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQTSSESSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILRNVRQSLVGLGAPYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILRNVRQSLVGLGAPYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCPPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCPPL 496
QY 481 PRGQARWQCDVAVLYGEOQPGWRFGAALTIVLDVNGDKLTDTVAICAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLYGEOQPGWRFGAALTIVLDVNGDKLTDTVAICAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGACQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGACQGHVLLRSQ 616
QY 601 PVLVRKVAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 660
DB 617 PVLVRKVAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGTLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGTLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736

QY 721 SLVGTPLSAFQNLRPVLAEDAQRLLFTALFPFKKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFQNLRPVLAEDAQRLLFTALFPFKKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRSLACESASSTEV 840
DB 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRSLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
DB 857 SGALKSTSCSINHPIFFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916
QY 901 QLELPVYAVYVMVYTSYKLVNFNTASENTSRVMQHQVQVSNLQGRSLPISLVFLVPV 960
DB 917 QLELPVYAVYVMVYTSYKLVNFNTASENTSRVMQHQVQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1036
QY 1021 PFGIOBEFNATLKGNSLFDWYIKTSHNHLITVSTAEILFNDSVFTLLPGQGFVRSOTET 1080
DB 1037 PFGIOBEFNATLKGNSLFDWYIKTSHNHLITVSTAEILFNDSVFTLLPGQGFVRSOTET 1096
QY 1081 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGPFKRYQKDMMBEGGPPGAEPQ 1137
DB 1097 KVEPFEVNPPLIVGSSVGGLLLLALITAAALYKLGPFKRYQKDMMBEGGPPGAEPQ 1153

RESULT 4
AAU80252
ID AAU80252 standard; Protein; 1153 AA.
AC AAU80252;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human integrin 1 alpha-M subunit protein.
XX
KW Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Key Location/Qualifiers
FT Misc-difference 499..500
FT /note= "Encoded by GCG CAG AGG"
XX
XX WO200218583-A2.
XX
XX 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US27227.
XX
XX 01-SEP-2000; 2000US-229700P.
XX
XX (BLOO-) CENT BLOOD RES INC.
XX
XX Springer TA, Shimoaka M, Lu C;
XX
XX WPI: 2002-382964/41.
XX
XX N-PSDB; ABK50046.
XX
XX Modified integrin-I or integrin I-like domain polypeptide useful as an
XX immunogen to produce antibodies specific to polypeptide, comprises a
XX disulfide bond such that polypeptide is stabilized in a desired
XX conformation.
XX
XX Disclosure; Page 109-112; 112pp; English.
XX
XX This invention relates to a modified integrin-I or integrin I-like

CC domain polypeptide comprising at least one disulfide bond so that the
 CC domain is stabilised in a desired conformation. The polypeptide of
 CC the invention may have anti-inflammatory or immunosuppressive activities.
 CC The polypeptides of the invention have an open conformation and are
 CC useful as immunogens to produce antibodies that selectively bind to
 CC integrin I-domain; and for identifying a modulator of integrin activity,
 CC or of interaction of an integrin and a cognate ligand. The polypeptide
 CC of the invention, or antibodies (preferably anti-LFA-1 antibody)
 CC is useful for treating or preventing an integrin mediated disorder which
 CC is an inflammatory or autoimmune disorder in a subject and for
 CC inhibiting the binding of an integrin to a cognate ligand such as Crohn's
 CC disease, nephritis; human immunodeficiency virus (HIV), myocardial
 CC infarction, Sjorgen's syndrome, rheumatoid arthritis, dermatitis.
 CC A therapeutic composition comprising the peptide of the invention is
 CC useful for treating an integrin mediated disorder in a subject. The
 CC polypeptides and/or active or antigenic fragments are useful as
 CC reagents for diagnosis of integrin-mediated disorders. The present
 CC sequence represents the human integrin-1 alpha-M protein subunit used to
 CC generate the mutant polypeptides of the invention.
 XX
 SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 23; Length 1153;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDENAMTFQENARGFGQVVQLQSGRVVVGAPQEIIVAAANGRGLYQCDYSTGSCPEI 60
 DB 17 FNLDENAMTFQENARGFGQVVQLQSGRVVVGAPQEIIVAAANGRGLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVYKGLCFGLGNSLRQOPK 120
 DB 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVYKGLCFGLGNSLRQOPK 136
 QY 121 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQLKSKTLFSLMQVSEEF 180
 DB 137 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQLKSKTLFSLMQVSEEF 196
 QY 181 RIHFTFEFQNNPNPSLVKPIITQLLGRTHATGVKRVIRELLNITNGARKNAFKLILVI 240
 DB 197 RIHFTFEFQNNPNPSLVKPIITQLLGRTHATGIRKVRRELFNITNGARKNAFKLILVI 256
 QY 241 TDGEKEGDPGLGYEDVPIEADREGVIRYVIGVGDAFSEKSRQELNTIASKPRDRHFQVN 300
 DB 257 TDGEKEGDPGLGYEDVPIEADREGVIRYVIGVGDAFSEKSRQELNTIASKPRDRHFQVN 316
 QY 301 NFEALKTIONQUREKIFAIEGTQTGSSSFHEHMSQEGFSAAITSGNPLLLSTVGSYDWAG 360
 DB 317 NFEALKTIONQUREKIFAIEGTQTGSSSFHEHMSQEGFSAAITSGNPLLLSTVGSYDWAG 376
 QY 361 GVPLYTSKEKSTFINNTRVDSQNDAYLGYAAAILLRNVOSIIVLGAPRYQHIGLVAMER 420
 DB 377 GVPLYTSKEKSTFINNTRVDSQNDAYLGYAAAILLRNVOSIIVLGAPRYQHIGLVAMER 436
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQSVCP 480
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQSVCP 496
 QY 481 PRQQRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTVDVAIGAPGEEDNRGAVYLF 540
 DB 497 PRQQRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTVDVAIGAPGEEDNRGAVYLF 556
 QY 541 HGTSGSGISPSHSORITAGSKLSPRLOYFGQSLSGGGDLTMDGLVDLTGVAQGHVLLRSQ 600
 DB 557 HGTSGSGISPSHSORITAGSKLSPRLOYFGQSLSGGGDLTMDGLVDLTGVAQGHVLLRSQ 616
 QY 601 PVLVRKAIMEFNEPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
 DB 617 PVLVRKAIMEFNEPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
 QY 661 YDLALDSGRPHSRPAVFNETKNSRRQTVQLGLTQTCTETLKLQLPNCIEDPVPSPVLRNLF 720
 DB 677 YDLALDSGRPHSRPAVFNETKNSRRQTVQLGLTQTCTETLKLQLPNCIEDPVPSPVLRNLF 736

QY 721 SLVGTPLSAFGNLRPVLAEADQRLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSAFGNLRPVLAEADQRLFTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRORSRWLACESASSTEV 840
 DB 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRORSRWLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPRTNKTEF 900
 DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPRTNKTEF 916
 QY 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 960
 DB 917 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 976
 QY 961 RLNQTVIWDPRPOVTVSENLSSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
 DB 977 RLNQTVIWDPRPOVTVSENLSSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
 QY 1021 PFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEIILFNDSVFLLPQCGAFVRSOTET 1080
 DB 1037 PFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEIILFNDSVFLLPQCGAFVRSOTET 1096
 QY 1081 KVEPEFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1137
 DB 1097 KVEPEFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1153

RESULT 5

AAO14428
 ID AAO14428 standard; protein; 1153 AA.
 XX AC AAO14428;
 XX DT 03-MAY-2002 (first entry)
 XX DE Integrin Mac-1 alpha subunit.
 XX KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
 KW open conformation; integrin related inflammatory disorder;
 KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
 KW reperfusion; hypovolemic shock; infarction; cerebral shock;
 KW viral infection; cancer; gene therapy; vaccine;
 KW bioactive agent screening.
 XX OS Unidentified.
 XX PN WO200204521-A2.
 XX PD 17-JAN-2002.
 XX PF 09-JUL-2001; 2001WO-US21805.
 XX PR 07-JUL-2000; 2000US-216600P.
 XX PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 XX PI (BLOO-) CENT BLOOD RES.
 XX PI Springer T;
 XX WP1; 2002-148167/19.
 XX PT New integrin I domain protein having alteration in at least 2
 PT noncontiguous regions and exits in an open conformation, useful for
 PT treating, preventing or suppressing inflammatory or immunological
 PT disorders
 XX PS Example 1; Fig 1F; 90pp; English.
 XX CC The invention comprises structurally biased variant integrin inserted (I)
 CC domain proteins, wherein the alterations to the protein occur in at least

FT Modified-site 1045..1047
FT /*label= putative N-glycosylation site
FT Modified-site 1051..1053
FT /*label= putative N-glycosylation site
FT Modified-site 1076..1078
FT /*label= putative N-glycosylation site
FT region 1..16
FT /*label= signal peptide
FT region 1106..1134
FT /*label= putative transmembrane region
XX
PN EP364690-A.
XX
XX 25-APR-1990.
XX
XX 17-AUG-1989; 89EP-0115159.
XX
XX 23-AUG-1988; 88US-0235353.
PR 09-MAR-1989; 89US-0321239.
XX
XX (DAND) DANA FARBBER CANCER INST INC.
XX
XX Springer TA, Corbi A;
XX WPI; 1990-125938/17.
DR N-PSDB; AAQ04043.
XX
XX New pure Mac-1 alpha sub-unit and corresp. DNA useful for treating
FT inflammation and viral infections, and in diagnosis
FT
XX Disclosure; Page ?; ?pp; English.
XX
XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
CC recognition of and migration to sites of inflammation. It also attaches
CC to cellular substrates as part of this function making it useful in
CC visualising endothelial tissue.
CC Mac-1 is a member of the Integrin Gene superfamily.
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 1153 AA;

Query Match 99.7%; Score 5858; DB 11; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1132; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQETVAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQETVAANQRGSLYQCDYSTGSCPEI 76
Qy 61 RLOVPVEAVNMSLGLSLAATSPQALLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 120
Db 77 RLOVPVEAVNMSLGLSLAATSPQALLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGIIIPHDPRMKEFVSTVMEQKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGIIIPHDPRMKEFVSTVMEQKSKTLFSLMQYSEEP 196
Qy 181 RIHFTFEFQNNPRSLVXPIQLLGRTHATGVRKVIKRELTNGARKNAFKILVI 240
Db 197 RIHFTFEFQNNPRSLVXPIQLLGRTHATGVRKVIKRELTNGARKNAFKILVI 256
Qy 241 TDGEKFGDPLGYEDVIPAEDREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRPHVFOVN 300
Db 257 TDGEKFGDPLGYEDVIPAEDREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRPHVFOVN 316
Qy 301 NFEALKTIONOLREKIPIAETGOTGSSSSFEHMSQEGFSAATNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONOLREKIPIAETGOTGSSSSFEHMSQEGFSAATNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFFINMTVDSMDNDAYLGYAAAIIILNRNQSVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFFINMTVDSMDNDAYLGYAAAIIILNRNQSVLGAPRYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGAGSLCSVDVDSNGSGLDVLIGAPHYYEOTRGQSVSCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGAGSLCSVDVDSNGSGLDVLIGAPHYYEOTRGQSVSCPL 496
Qy 481 PRGORARWOCDAVLYGEGQGPWRFGAALTVLGDVNGDKLTDAIGAAGEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLYGEGQGPWRFGAALTVLGDVNGDKLTDAIGAAGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORJAGSKLSPRLQYFGOSLSSGGDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORJAGSKLSPRLQYFGOSLSSGGDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLVRKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIOSVVT 660
Db 617 PVLVRKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIOSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRVAVNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 840
Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQOVSNLQORSLSISLVLVLPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQOVSNLQORSLSISLVLVLPV 976
Qy 961 RLMQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1020
Db 977 RLMQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1036
Qy 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDSVFTLLPGQGAFVRSQTE 1080
Db 1037 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDSVFTLLPGQGAFVRSQTE 1096
Qy 1081 KVEPPEVPNPLPIVGVSSVGLLALLITAAALYKLGFFKQYKDMNSEGGPPGAEPQ 1137
Db 1097 KVEPPEVPNPLPIVGVSSVGLLALLITAAALYKLGFFKQYKDMNSEGGPPGAEPQ 1153
RESULT 7
AAR07120
ID AAR07120 standard; protein; 1163 AA.
XX
AC AAR07120;
XX AC
XX 25-MAR-2003 (updated)
DT 05-FEB-1991 (first entry)
XX
XX p150.95 alpha subunit encoded by clone lambdaX47.
DE
XX p150.95 leucocyte adhesion receptor alpha-subunit;
KW hairy cell leukaemia; rhinovirus.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 61..63
FT /label-glycosylation site
FT Modified-site 89..91
FT /label-glycosylation site
FT Modified-site 385..387
FT /label-glycosylation site

FT Modified-site 392..394 /label=glycosylation site
 FT Modified-site 697..699 /label=glycosylation site
 FT Modified-site 735..737 /label=glycosylation site
 FT Modified-site 899..901 /label=glycosylation site
 FT Modified-site 904..906 /label=glycosylation site
 FT Modified-site 939..941 /label=glycosylation site
 FT Modified-site 1050..1052 /label=glycosylation site
 FT Domain 1108..1133 /label=transmembrane
 FT Region 1..19 /label=signal peptide
 FT Region 20..44 /label=N-terminus

XX WO9010646-A.
 PN 20-SEP-1990.
 PD 09-MAR-1990; 90WO-US01257.
 XX 02-FEB-1989; 89US-0305458.
 PR (DAND) DANA FARBER CANCER INST INC.
 PA Corbi AA, Springer TA;
 PI WPI; 1990-304985/40.
 DR N-PSDB; AAQ06068.
 XX Treatment of viral esp. rhino-viral infection - by admin. of alpha
 PT sub-unit of p150.95 cell surface adhesion receptor, opt. together
 PT with a beta chain of CD-18 family.
 XX Disclosure; Fig 3; 59pp; English.
 PS Clone lambda X47 was isolated from a cDNA library constructed from
 CC total RNA extracted from phorbol myristate acetate stimulated HL-60
 CC myelomonocytic cells. The library was screened with oligonucleotide
 CC probes based on tryptic peptide fragments of p150.95. The sequence
 CC can be attached to appropriate control elements and expressed in
 CC prokaryotic and eukaryotic cells. The protein can be used to treat
 CC or prevent rhinoviral infection because it interacts with ICAM-1
 CC and inhibits cell-virus attachment. It can also be used as an
 CC anti-inflammatory agent.
 CC See also AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 1163 AA;

Query Match 59.0%; Score 3464; DB 11; Length 1163;
 Best Local Similarity 61.1%; Pred. No 3e-279;
 Matches 690; Conservative 138; Mismatches 295; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCBPI 60
 DB 20 FNLDTEELTAFRVDSAGFSDSVVQVANSWVVGAPQKITAANQTGGLYQCGYSTGACBPI 79
 QY 61 RLQVPVEANMISLGLSLAATSPOLLACGPTVHQCSTENTYVKGCLFLGSLNLRQPOK 120
 DB 80 GLQVPPEANMISLGLSLAATSPOLLACGPTVHQCSTENTYVKGCLFLGSLNLRQPOK 137
 QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHDIFRAMEKFEVSTVMEQLKKSKTLFSLMOYSEF 180
 DB 138 LPVSRQECPRQSDIVFLDGGSGISRRNFATMMNFVRAVISQFORPSTQFSLMQFSNKF 197

QY 181 RIHFTKEFQNNPNBSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
 DB 198 QTHFTFEFRNTSNPLSLASVHQLQGFYITATQVNVHRLPHASVGAARDATKILIVI 257
 QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYIGVDAFRSEKSRQBLNTIASKPPDRDHVQVN 300
 DB 258 TDGKKGDSLKYKDVIPMAADAAGIIRYAGVGLAFQNRNSWKLNDIASKPSQEHFQVE 317
 QY 301 NFEALKTIONLREKIPAIETGTGTGSSSSFEHEMSQEGFSAATISNGPLLSVTSVGDWAG 360
 DB 318 DFDALDKIQNLKEKIPAIETGTGTGSSSSFEHEMSQEGFSAATISNGPLLSVTSVGDWAG 377
 QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAAILRNRYQSLVGLAPRYQHIGLVAMFR 420
 DB 378 GAFLYPNWSPPTFINNSQENVMDRDSYLGYSTELALWGVQSLVGLAPRYQHTGKAVIPT 437
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQSVCP 480
 DB 438 QVSRQWRMAEVTGTQIGSYFGASLCSVDVDTGSTDVLVIGAPHYEYQTRGGQSVCP 497
 QY 481 PRGORARWQCDVLYGEOGQPGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 540
 DB 498 PRGWR-RWMCDAVLYGEOGQPGWGRFGAALTIVLGDVNGDKLTDVWIGAPGEENRGAYLYF 556
 QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
 DB 557 HGVLPSPISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
 QY 601 PVLRVKAIMBPNPREVARNVFECDQVYVKEAGEVRVCLVQKSTRDLREGIQSVVT 660
 DB 617 PVLWVGVSQMFIPAEIPRSFAFECREQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
 DB 677 LDALDPGLSPRATFOETKNSRVRVLGLKAHCENFNLLPSCVEDSVTPTLRNF 736
 QY 721 SLVCTPLSACGNLRPLAEDAQRLFTALFPKCNCGNDNICODDLSITFSFMSDCLVVG 780
 DB 737 TLVGKPLLAFLNLRPLAEDAQRYFTASLPKCNCGADHICQDNLGISFSPGLKSLVG 796
 QY 781 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 DB 797 SNLELNAEVMVNDGSDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 854
 QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTPE 900
 DB 855 SQGTWSTSCRINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTPE 914
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTAS-ENTSRVMQHOYQVNSLQORSUPISLVFLVP 959
 DB 915 QLELPVKYAVYVMTSHGVSTKYLNFTAS-ENTSRVMQHOYQVNSLQORSUPISLVFLVP 974
 QY 960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPFSDHDLAELKAPVAVNCISAVCORIQCDI 1019
 DB 975 VELNQAEMVMDVEVSHPNPSLRCSQKIAPASDFLAHIQKNPVLDCSAGLFRCDV 1034
 QY 1020 PFGIQGEFNATLKNLSFDWYIKTSNNHLLIVSTAELFNDVSFTLLPGGAFVRSQTE 1079
 DB 1035 PSFSVQBELDPTLKNLSFGWVRQILQKVSWSVAEITFDTSVYSQLPQGERFMRAQTT 1094
 QY 1080 TKVEPPEVNPPLIVGVSSVGLLLALITAAALYKLGFFKQYKDMSE 1128
 DB 1095 TVLEKYKVNPTPLIVGVSSVGLLLALITAAALYKLGFFKQYKDMSE 1143

RESULT 8

AAW65091
 ID AAW65091 standard; Protein; 1163 AA.

XX AC AAW65091;

XX DT 28-SEP-1998 (first entry)

XX

OS Homo sapiens.
PN WO200029446-A1.
XX 25-MAY-2000.
XX 16-NOV-1999; 99WO-US27139.
XX 16-NOV-1998; 98US-0193043.
PR 08-JUL-1999; 99US-0350259.
XX (ICOS-) ICOS CORP.
XX Gallatin MW, Van Der Vieren M;
XX WPI; 2000-387751/33.
DR
XX
XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous
PT system injury sites -
XX
PS Example 5; Fig 1; 270pp; English.
XX
CC Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11c. This
CC sequence was used in an alignment to identify a novel beta2 integrin
CC alpha subunit: alpha_d (AA060014 and AAB07359). The present sequence has
CC approximately 66% identity to the protein sequence of alpha_d. The
CC Alpha_d gene and protein may be useful in therapy for diseases linked
CC to aberrant alpha_d function e.g. Type I diabetes, atherosclerosis,
CC multiple sclerosis, asthma, psoriasis, lung inflammation, acute
CC respiratory distress syndrome, rheumatoid arthritis and leukocyte
CC adhesion deficiency (LAD). In addition, anti-alpha_d monoclonal
CC antibodies may be used in the inhibition of macrophage infiltration at
CC the site of a central nervous system injury. The monoclonal antibodies
CC can also be used to detect and diagnose Crohn's disease.
XX
SQ Sequence 1163 AA;
Query Match 58.7%; Score 3450; DB 21; Length 1163;
Best Local Similarity 61.0%; Pred. No. 4.4e-278;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;
QY 1 FNLDTENAMTFQENARFGQSVVLOGSRVVGAPQEIIVAAQORGLSYOCDYSTGSCEPI 60
DB 20 FNLDTEELTAFRVDSAGFQSVVQYANSWVVGAPQKIIAAQIGLYOCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFLGSLNLRQPOK 120
DB 80 GLQVPEAVNMSLGLSLASTTSQQLACGPTVHCEGRNMYLTGLCFLLGPT--QLTQR 137
QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHPFRMKFVSTVMEQLKKSKTLFSLMQVSEEF 180
DB 138 LPVSRQECPRQEDIVFLDGGSGISSRNEFATMNFVRAVISQFQRPSTQFSLMQFSNKF 197
QY 181 RHIFTEKFQNNPNRSLVKPIITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
DB 198 QTHFTPEEPRTNPLSLASVHLQLOGFTYTATAIGNVHRLFHASVGAARRDAIKILIVI 257
QY 241 TDEKFGDPLGYEDVPEADREGVIRVIGVGAFAKSEKRSQELNLTIASPPRDHVFQVN 300
DB 258 TDGKKGSDLDYKDVIPMAADAGIIRVAGVLAQFNRSWKELNDIASKPSQEHFKVE 317
QY 301 NFALKTIQNLREKIPAEICTGTGSSSSEHEMSQEGFSAAITSNGLPSTLVGSDVWAG 360
DB 318 DFDALDKIQNLREKIPAEIGTETISSSSSELEMAQEGFSAVFTPDGPVLGAVGSEFTWSG 377
QY 361 GVFLYTSKESKSTFINMTNRVDSNDNDAYLGVAAILLRNVQSLVLGAPRYQHIGLVAMFR 420

DB 378 GAFLYPPNMSPTPINMSQENVDMRDSYLGYSYSELALWKQSVLGAAPRYQHICKAVIFI 437
QY 421 QNTGHWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGABHYEYQTRGGQSVCP 480
DB 438 QVSRQWRMAEVIQIGSYFGASLCSVDVDVDTGSDTLVLIGABHYEYQTRGGQSVCP 497
QY 481 PRGORARWOCDAVLYGEOGPWGRFGAALTVLGVDVNGDKLTDVAIGAPEGEDNRGAVYLF 540
DB 498 PRGWR-RWDCDAVLYGEOGPWGRFGAALTVLGVDVNGDKLTDVIGAPGEENRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDQTLMDGLVDLTVAAGHVLRLRSQ 600
DB 557 HGVLGPSISPSHSQRIAGSKLSPRLQYFGQSLSGQDQTLMDGLVDLTVAAGHVLRLRSQ 616
QY 601 PVLRYKALMEFNPREVARNFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLWVGSMQPIPAEIPRSAFECEQVQVSEQLVQSNICLYIDKRSKNLGLSRLDQSSVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
DB 677 LDALAPGRLSFRAIFQETKRSLSRVRLGLKAHCENFNLLSCVEDSVIPILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALFPFEKNCNDNICQDDLSITFSFMSDCLVWG 780
DB 737 TLVCKPLLAFLNRLPMLAALAQRYFTASLPFEKNCADHICQDNLGISFSFGLKSLVWG 796
QY 781 GPRFNVTVVRNDEGSYRTQVTFPPFLDLSYKVTSLQNRQSRWRLACESASSTEV 840
DB 797 SNLELNAEVMVMNDGDSYGTITTFSPAGLSYRYVAEGQKQQLRSUHLTC--CSAPVG 854
QY 841 SCALKSTSCSNHPIFFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
DB 855 SQGTWSTSCRNLHIFRGAQITFLATPDVSPKAVGLDRLLLIANVSSENNIPRISKTF 914
QY 901 QLELPKYAVVMVTSHGVTSKYLNFTAS-ENTSEVMQHYQVSNLQORSLPISLVFLVP 959
DB 915 QLELPKYAVIVVSSHQFTKYLNFSESEKESHVAMHYQVNNLQORDLPVSNFVWP 974
QY 960 VRLNQTIVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCISAVCQICDI 1019
DB 975 VELNQEAVMQVDEVSHPNQPSLRCSESEKIAPPASDFLAHIQKNPVLDCSIAGCLFRCDV 1034
QY 1020 PFGIQTQEFNATLKNLSFDWYIKTSHNHLIVSTAEILVNDSTVFTLLPGQCAFVRSOTE 1079
DB 1035 PSFSVQEEUDFTLKNLSFGWVRQLQKVSVSVVAEIIFTSVYSQLPQGEAFWRAQTI 1094
QY 1080 TKVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSE 1128
DB 1095 TVLEKYKVNPIPLIVGSSIGGLLLALITAVLYKVGFFKQYKEMEE 1143
RESULT 10
ABG61470
ID ABG61470 standard; Protein; 1163 AA.
XX ABG61470;
XX
XX 27-AUG-2002 (first entry)
XX Human Beta2 integrin alphaCD11c subunit.
XX
XX Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit;
KW LAD; leukocyte adhesion deficiency; inflammatory response; diabetes;
KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;
KW intracellular cell adhesion molecule; vascular cell adhesion molecule;
KW locomotor recovery; locomotor damage; locomotor impairment;
KW autonomic dysfunction; sensory dysfunction; spinal cord injury.
XX
OS Homo sapiens.
XX

PF 08-APR-2002; 2002WO-US10824.
XX
PR 06-APR-2001; 2001US-281731P.
PR 06-APR-2001; 2001US-281732P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Jay G;
XX
DR WPI; 2003-058520/05.
XX
XX Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
PT patients -
XX
PS Claim 1; Page 225-228; 416pp; English.
XX
CC The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.
XX
SQ Sequence 1163 AA;
Query Match 58.7%; Score 3448; DB 24; Length 1163;
Best Local Similarity 60.9%; Pred. No. 6.5e-278;
Matches 688; Conservative 136; Mismatches 299; Indels 6; Gaps 4;
1 ENLDTENAMTFOENARGFGQSVVVLQGGSRVVVGAPOEIVAAQNRGLYOCYSTGSCPEI 60
20 FNLDTTELTAFRVDSAGFGDSVVQVANSVVVGAPOKITAANQTGGLYOCYSTGACEPI 79
61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVVKGLCLFLGSLNRQOQPK 120
80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRNMYLTGLCLLGLPT--QLTOR 137
121 FREALRGQOQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKKSKTLFSLMQVSEEF 180
138 LPVSRQECRQEQDIIVFLIDGSGSISSRNFAFMNFVRAVISQFORPSTQFSLMOFSNKF 197
181 RIHFTKEFQNNPNRPSLVKPIITOLLGRTHATGVKVIPELLNITNGARKNAFKILIVI 240

Db 198 QTHLTPEPRRTSNPLSLASVHLOQGYTATATAIQNVVHRLPHASYGARRDATKILIVI 257
Qy 241 TDEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSEKROELINTIASPPRDHVPQVN 300
Db 258 TDGKKGDTLDYKDVIPMDAAGIIRIYAGVLAQNRNSWKELNDJASKPSOEHIFKVE 317
Qy 301 NFEALKTIQOLREKIFAIEGTOTGSSSSPEHEMSOEGFSAAITNSGPLLSITVGSVDWAG 360
Db 318 DFDAKDIQOLREKIFPIEGTETSSSFELEMAQEGFSAVFTPDGVLGAVGSPWMSG 377
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNRVOSLVLAGAPRYOHIGLVAMPR 420
Db 378 GAFLYPPNMPSTFINMSQENVMDRDSYLGYSTELALWKGVQSLVLAGAPRYOHTGKAVFT 437
Qy 421 QNTGHWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQVSCPL 480
Db 438 QVSRQRMKAEBVTGTQIGSYFGPSLCSVDVDSGSDTLVLIGPHYVEQTRGAQVSCPL 497
Qy 481 PRGORARWQCDVLYGEGQOPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLFF 540
Db 498 PRGWR-RWNCDAVLYGEGQHPWGRFGAALTVLGDVNGDKLTDVIGAPGEEHNGAYLFF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLPSPISPSHSQRIAGSLSRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLLR 616
Qy 601 PVLRYKAIMFENPREVARNVFECDQVYVKGKEAGEVVRVCLHVQKSTRDRLEGOIQSVWT 660
Db 617 PVLWVGSMQFIPAEIPRSAFECEQVVSQTLVQSNICLYIDKRSKNSLGRDLSQSVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRTOVLGLTCTETLKLQLPNCIEDPVSPTVLRNLF 720
Db 677 LDLDLDFGRLSPRATFQETKNRSVRVVLGAKHCENFNLPLPSCVEDSVPTILNLP 736
Qy 721 SLVGTPLSAGFNRLPVLAEADAQRLFTALFPFKNCGNDNICQDDLSITFSFMSLCLVWG 780
Db 737 TLVCKPLLPNLAPMLAADAQRYFTASLPFKNCGADHICQDNLGIFSFPGLKSLVG 796
Qy 781 GPRFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSRLACESASSTEV 840
Db 797 SNLELNAEVMVMVNDGDSYGTITTFSPAGLSYRYVAEGQKQGLRSLHLTCDSPAVG-- 854
Qy 841 SGALKSTSCSINHPIFFENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTFF 900
Db 855 SQGTWSTSCRINHILFRGGAQITFLATFDVSPKAVLGDRLLLLTANVSENNTPRTSKTF 914
Qy 901 QLELPVKYAVVMVYVTSKYLNFNTAS-ENTSRVMQHOYQVSNLQORSUPISLVFLVP 959
Db 915 QLELPVKYAVVMVYVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFVWP 974
Qy 960 VRLNQTVIWDPRQVTFSENLSSTCHTKEPLPSHSDFLAELRKA PVNVCIAVCQRIQCDI 1019
Db 975 VELNQEAVMMDVEVSLPQNPSLRCSEKIAQIPASDFLAHIQKNPVLDCSIAGLRFRCDV 1034
Qy 1020 PFFGIQEBFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSVFTLLPGCAFVRSQTE 1079
Db 1035 PSFSVQEBLDTFLKGNLSFGVWRQILQKVSQVSVAEITFTDSYVSQLPGQEAFWRAQT 1094
Qy 1080 TKVPEFPEPNPLPIVIGSSVGLLLALITAAALYKLGFFKQYKDMMS 1128
Db 1095 TVLEKYKVHNPTPLIVGSSIGGLLLALITAVLYKVGFQKQYKEMMEE 1143
RESULT 12
AAR78166
ID AAR78166 standard; Protein; 1161 AA.
XX
AC AAR78166;
XX
DT 28-DEC-1995 (first entry)
XX
DE Human beta-2 integrin alpha-d.
XX

[illegible]

multiple sclerosis; asthma; psoriasis; lung inflammation; acute respiratory distress syndrome; rheumatoid arthritis.

Homo sapiens.

Key	Location/Qualifiers
1	1.1
2	2.1
3	3.1
4	4.1
5	5.1
6	6.1
7	7.1
8	8.1
9	9.1
10	10.1
11	11.1
12	12.1
13	13.1
14	14.1
15	15.1
16	16.1
17	17.1
18	18.1
19	19.1
20	20.1
21	21.1
22	22.1
23	23.1
24	24.1
25	25.1
26	26.1
27	27.1
28	28.1
29	29.1
30	30.1
31	31.1
32	32.1
33	33.1
34	34.1
35	35.1
36	36.1
37	37.1
38	38.1
39	39.1
40	40.1
41	41.1
42	42.1
43	43.1
44	44.1
45	45.1
46	46.1
47	47.1
48	48.1
49	49.1
50	50.1
51	51.1
52	52.1
53	53.1
54	54.1
55	55.1
56	56.1
57	57.1
58	58.1
59	59.1
60	60.1
61	61.1
62	62.1
63	63.1
64	64.1
65	65.1
66	66.1
67	67.1
68	68.1
69	69.1
70	70.1
71	71.1
72	72.1
73	73.1
74	74.1
75	75.1
76	76.1
77	77.1
78	78.1
79	79.1
80	80.1
81	81.1
82	82.1
83	83.1
84	84.1
85	85.1
86	86.1
87	87.1
88	88.1
89	89.1
90	90.1
91	91.1
92	92.1
93	93.1
94	94.1
95	95.1
96	96.1
97	97.1
98	98.1
99	99.1
100	100.1

Key	Location, Year
Domain	17..1108

```
/label= Extracellular_domain
```

```
Domain;
1109..1128
/label= Transmembrane domain
/note= "homologous to the human CD11c transmembrane
region"
```

Domain	region
1129..1161	

/label= Cytoplasmic domain

Domain: /note= "region homologous to the I (insertion) domain common to Cplla, Cpllb and Cpllc"

WO9731099-A1.

28-AUG-1997.

24-FEB-1997: 97WO-US02713.

22-FEB-1996: 96U

(ICOS-) ICOS CORP.

Gallatin WM. Van De

WPI: 1997-435154/

Hybridoma 199M and antibody secreted by it - specific for new rat beta2: integrin subunit, useful to detect subunit in cells and modulate its activity

Example 5; Page 116-120; 222pp; English.

This polypeptide comprises a novel human beta 2 integrin subunit, designated alpha d. Its sequence was deduced from a cDNA clone (see A4779220) isolated from a spleen cDNA library. Alpha d is involved in cell migration, phagocytosis and cell-cell interaction. Recombinant alpha d polypeptides can be expressed in transformed host cells for use in assays for identifying antibodies or other compounds that modulate alpha d activity or which modulate the interaction between alpha d and a ligand, for treating or preventing diseases in which macrophages are implicated. Treatment is applicable to disease states in which alpha d binding, or localised accumulation of cells which express alpha d, is implicated such as such as type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome and rheumatoid arthritis.

Sequence 1161 AA;

Very Match	58.24	Score 3417	DB 18	Length 1161
------------	-------	------------	-------	-------------

Model	Best Local Similarity	Pred. No. 2.5e-275;
Model 1	59.8%	
Model 2	59.8%	
Model 3	59.8%	
Model 4	59.8%	
Model 5	59.8%	
Model 6	59.8%	
Model 7	59.8%	
Model 8	59.8%	
Model 9	59.8%	
Model 10	59.8%	
Model 11	59.8%	
Model 12	59.8%	
Model 13	59.8%	
Model 14	59.8%	
Model 15	59.8%	
Model 16	59.8%	
Model 17	59.8%	
Model 18	59.8%	
Model 19	59.8%	
Model 20	59.8%	
Model 21	59.8%	
Model 22	59.8%	
Model 23	59.8%	
Model 24	59.8%	
Model 25	59.8%	
Model 26	59.8%	
Model 27	59.8%	
Model 28	59.8%	
Model 29	59.8%	
Model 30	59.8%	
Model 31	59.8%	
Model 32	59.8%	
Model 33	59.8%	
Model 34	59.8%	
Model 35	59.8%	
Model 36	59.8%	
Model 37	59.8%	
Model 38	59.8%	
Model 39	59.8%	
Model 40	59.8%	
Model 41	59.8%	
Model 42	59.8%	
Model 43	59.8%	
Model 44	59.8%	
Model 45	59.8%	
Model 46	59.8%	
Model 47	59.8%	
Model 48	59.8%	
Model 49	59.8%	
Model 50	59.8%	
Model 51	59.8%	
Model 52	59.8%	
Model 53	59.8%	
Model 54	59.8%	
Model 55	59.8%	
Model 56	59.8%	
Model 57	59.8%	
Model 58	59.8%	
Model 59	59.8%	
Model 60	59.8%	
Model 61	59.8%	
Model 62	59.8%	
Model 63	59.8%	
Model 64	59.8%	
Model 65	59.8%	
Model 66	59.8%	
Model 67	59.8%	
Model 68	59.8%	
Model 69	59.8%	
Model 70	59.8%	
Model 71	59.8%	
Model 72	59.8%	
Model 73	59.8%	
Model 74	59.8%	
Model 75	59.8%	
Model 76	59.8%	
Model 77	59.8%	
Model 78	59.8%	
Model 79	59.8%	
Model 80	59.8%	
Model 81	59.8%	
Model 82	59.8%	
Model 83	59.8%	
Model 84	59.8%	
Model 85	59.8%	
Model 86	59.8%	
Model 87	59.8%	
Model 88	59.8%	
Model 89	59.8%	
Model 90	59.8%	
Model 91	59.8%	
Model 92	59.8%	
Model 93	59.8%	
Model 94	59.8%	
Model 95	59.8%	
Model 96	59.8%	
Model 97	59.8%	
Model 98	59.8%	
Model 99	59.8%	
Model 100	59.8%	

atches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

1 FNLDTENAMTFQENARGFGQSVVLQGSRVVVGAPQEIVAANQRGSLYQCDYSTGSCEPI 60

17 FNLDVEEPTIFQEDAGGFGQSVQFGGSRLLVVGAPLEVVAAANQTGRRLYDCAAATGMCQPI 76

61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQTCSENTYVKGLCFLFGSNLRQPPQK 120

77 PLHIRPEAVNMSGLTLAASTNGSRLLACGPTLHRVCGENSYSKGSCLLGSRW-EIIQT 135

121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMQYSEEF 180

136 VPDATPECPHOEMD|VFL|D|GSGSI|DONDFNOMKGFVQAVMGOFEGTDTL|FALMQYSNLL 195.

181' RIHFTFEFONPNPRSLVKPITOLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240

[illegible]

RESULT 14

AAW72825

ID AAW72825 standard; Protein; 1161 AA.

AC AAW72825;

XX

DT 19-JAN-19

XX
XX

DE Human alpha-d.:

XX KW Alpha-d; beta-2 integrin alpha subunit; extracellular alpha-d;
 KW immunoglobulin; fusion protein; binding molecule; antibody;
 KW immunohistochemical analysis; diabetes; atherosclerosis; asthma;
 KW multiple sclerosis; psoriasis; lung inflammation; arthritis;
 KW acute respiratory distress syndrome.
 XX OS Homo sapiens.
 XX PN US9831029-A.
 XX PD 03-NOV-1998.
 XX PF 07-JUN-1995; 95US-0482293.
 XX PR 07-JUN-1995; 95US-0482293.
 XX PR 23-DEC-1993; 93US-0173497.
 XX PR 05-AUG-1994; 94US-0286889.
 XX PR 21-DEC-1994; 94US-0362652.
 XX PA (ICOS-) ICOS CORP.
 XX PI Gallatin WM, Van Der Vieren M;
 XX WPI, 1998-609318/51.
 XX DR N-PSDB; AAV67281.
 XX PT Antibodies specific for beta-2 integrin alpha-subunit d - useful in
 XX immunohistochemical analysis
 XX Example 5; Column 61-66; 106pp; English.
 XX The present sequence represents human alpha d. The present invention
 CC describes: (1) an antibody that specifically binds alpha d, which
 CC is a beta 2 integrin alpha-subunit polypeptide; (2) a monoclonal
 CC antibody as in (1); (3) an anti-idiotypic antibody specific for the
 CC monoclonal antibody of (2); (4) a hybridoma cell line producing the
 CC monoclonal antibody of (2). Antibodies specific for alpha d can be
 CC used in immunohistochemical analysis to localise alpha d to subcellular
 CC compartments or individual cells within tissues. Substances that modulate
 CC alpha d binding (which may include antibodies) can be used to treat
 CC diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
 CC inflammation, acute respiratory distress syndrome or arthritis.
 XX SQ Sequence 1161 AA;
 Query Match 58.2%; Score 3417; DB 19; Length 1161;
 Best Local Similarity 59.8%; Pred. No. 2.5e-275;
 Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;
 QY 1 FNLDTENAMTQENARGQSVVQLQGRVVGAPQETVAANQSGSLVQCDYSTGSCPEI 60
 DB 17 FNLDVEEPTIQEDAGGFGQSVVQFGGRLVVGAPLEVAANQGRLYDCAANTGMCQPI 76
 QY 61 RLQVPEAVNMVSLGLSLAATSPQLLACGPTVHTQTCSENTYVKGCLFLFGSLNRQOPQK 120
 DB 77 PLHIRPEAVNMVSLGLTAASTNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-ETIQT 135
 QY 121 FPEALRGCPQSDIAFLIDGSGIIPHDPFRMKEFVSTVMEQLKSKTLFSLMQYSEEP 180
 DB 136 VPDATPECPHQMIDIVFLIDGSGIDQDNDFQNMKGFOVQAVMGQPEGDTLTPALMQYNLL 195
 QY 181 RIHTEFEFQNNPNSRLVKPIITOLGRTHTATGVRKVIKRELLNITNGARKNAKILIVI 240
 DB 196 KIHTFTQFRYSPSQSLSDPVLQKGLTFTATGLTIVTQLFHHKNGARKSAXKILIVI 255
 QY 241 TDGKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN 300
 DB 256 TDGQKYKDPLEYSVIPAEXKAGIRYAIGVGHAFQGTARQELNTISSAPPQDHVKVD 315
 QY 301 NFEALKTTQNLQREKIFAETQRTGSSSSPHEHMSQEGFSAAITSNGLLSTVGSYDWAG 360
 DB 316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMWDLFLGAVGSFSWSG 375

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNRRVQSLVGLGAPRYQHIGLVAMFR 420
 DB 376 GAFLYPPNMSPTFINMSQENVDRDSYLGYSTELAWKGVQNLVGLGAPRYQHTGKAVIPT 435
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVILGAPHYHYEQTRGGQSVCP 480
 DB 436 QVSRQWRKKAETGTQIGSYFGASLCSVDVDSNGSTDLVILGAPHYHYEQTRGGQSVCP 495
 QY 481 PRGORARWQCDVLYGEGOPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYLYF 540
 DB 496 PRGORVQWQCDVLYRGEQGHGPRFGAALTVDGVNEDKLDVIGAPGQENRGAYLYF 555
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVDLTGVAQGHVLLRSQ 600
 DB 556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGQDLTODGLMDLAVGARGQVLLRLSL 615
 QY 601 PVLRVKAIMBENPREVARNVFECNDQVVKGEKAGEVRVCLHVQKSTDRDRREGIOISVVT 660
 DB 616 PVLKVGVMARFSPVEKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
 QY 661 YDLALDSGRPHSRAVFNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
 DB 674 FDLALDPGRUTSRAIFNETKNPTLTKRKTGLGLGHCTLKLPLDPCDEVVSPILHLNF 733
 QY 721 SLVGTPLSFAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCILVVG 780
 DB 734 SLVREPIPSQNLRLPVLAVGSQLPTASLPFEKNCQDGLCEGDLGVTLSFGSLQTLTVG 793
 QY 781 GPREFNVTVVRNGDEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 DB 794 SSLELNVITVMNAGDSYGTVSVLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852
 QY 841 SGALKSTCSINHPPIPEENSEVFNITFDVDSKASLGNKLLKXANVTSENNMPTNKTKEP 900
 DB 853 EG-LRSRCSVNHPIFHEGSGNGTFIVTFDVSYKATLDGRMLRASASNNKASSKATF 911
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNF-TASENTSRVQHQVQVSNLQORSIPSLVLVLP 959
 DB 912 QLELPVKYAVYVMTSHGVSTKYLNF-TASENTSRVQHQVQVSNLQORSIPSLVLVLP 971
 QY 960 VRLNQTVIMDRPQVTFSENLSSTCHTKERIPSHSDPLAELRKAPVNVCSIAVCORICDI 1019
 DB 972 VLLNGVAVWDMVMEAPSQSL--PCVSRKPPQHSDFLTQISRSPMLDCSIADCLQPRCDV 1029
 QY 1020 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIYVSTAEILFNDVSFTLLPGOGAFVRSQTE 1079
 DB 1030 PSFSVQEELOFTLKGNSLFGWVRETLQKKVLVSVSAEITFDTSVYSQLPQOEAFMQME 1099
 QY 1080 TKVEPPEFVNPPLPLIVGSSVGGILLALITAAALYKLGFFKRYQKDMSE 1128
 DB 1090 MVLEDEEVYNAIPIINGSSVGAULLALITATLYKLGFFKRYKEMLED 1138
 RESULT 15
 AAW65089
 ID AAW65089 standard; Protein; 1161 AA.
 XX AAW65089;
 AC AAW65089;
 XX 28-SEP-1998 (first entry)
 DT 28-SEP-1998 (first entry)
 XX Human Beta-integrin alpha-d subunit protein isolated from clone 19A2.
 DE Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome;
 KW rheumatoid arthritis.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Peptide 1..16
 FT

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 14:13:09 ; Search time 13.5618 Seconds
(without alignments)
3547.268 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGAEFQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5868	99.9	1153	1	US-08-173-497-3
2	5868	99.9	1153	1	US-08-286-889-3
3	5868	99.9	1153	1	US-08-485-618-3
4	5868	99.9	1153	1	US-08-362-652-3
5	5868	99.9	1153	2	US-08-605-672-3
6	5868	99.9	1153	2	US-08-482-293A-3
7	5868	99.9	1153	2	US-08-943-363-3
8	5868	99.9	1153	3	US-09-193-043-3
9	5868	99.9	1153	4	US-09-688-307A-3
10	5837.5	99.4	1152	2	US-08-476-062A-43
11	5837.5	99.4	1152	5	PCT-US96-01314-43
12	5837.5	99.4	1152	6	5424399-2
13	3473	59.1	1163	2	US-08-476-062A-44
14	3473	59.1	1163	5	PCT-US96-01314-44
15	3450	58.7	1163	1	US-08-173-497-4
16	3450	58.7	1163	1	US-08-286-889-4
17	3450	58.7	1163	1	US-08-485-618-4
18	3450	58.7	1163	1	US-08-362-652-4
19	3450	58.7	1163	2	US-08-605-672-4
20	3450	58.7	1163	2	US-08-482-293A-4
21	3450	58.7	1163	2	US-08-943-363-4
22	3450	58.7	1163	3	US-09-193-043-4
23	3450	58.7	1163	4	US-09-688-307A-4
24	3417	58.2	1161	1	US-08-173-497-2
25	3417	58.2	1161	1	US-08-286-889-2
26	3417	58.2	1161	1	US-08-485-618-2
27	3417	58.2	1161	1	US-08-362-652-2

28 3417 58.2 1161 2 US-08-605-672-2 Sequence 2, Appli
29 3417 58.2 1161 2 US-08-482-293A-2 Sequence 2, Appli
30 3417 58.2 1161 3 US-08-943-363-2 Sequence 2, Appli
31 3417 58.2 1161 3 US-09-193-043-2 Sequence 2, Appli
32 3417 58.2 1161 4 US-09-688-307A-2 Sequence 2, Appli
33 3401.5 57.9 1161 1 US-08-485-618-99 Sequence 99, Appli
34 3401.5 57.9 1161 2 US-08-605-672-99 Sequence 99, Appli
35 3401.5 57.9 1161 2 US-08-482-293A-99 Sequence 99, Appli
36 3401.5 57.9 1161 2 US-08-943-363-99 Sequence 99, Appli
37 3401.5 57.9 1161 3 US-09-193-043-99 Sequence 99, Appli
38 3401.5 57.9 1161 4 US-09-688-307A-99 Sequence 99, Appli
39 3239.5 55.1 1161 3 US-09-193-043-55 Sequence 55, Appli
40 3239.5 55.1 1161 4 US-09-688-307A-55 Sequence 55, Appli
41 3232.5 55.0 1161 1 US-08-485-618-55 Sequence 55, Appli
42 3232.5 55.0 1161 1 US-08-362-652-55 Sequence 55, Appli
43 3232.5 55.0 1161 2 US-08-605-672-55 Sequence 55, Appli
44 3232.5 55.0 1161 2 US-08-482-293A-55 Sequence 55, Appli
45 3232.5 55.0 1161 2 US-08-943-363-55 Sequence 55, Appli

ALIGNMENTS

RESULT 1

US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVVVGAPQBIIVANORGSLYQCDYSTGSGEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGRVVVVGAPQBIIVANORGSLYQCDYSTGSGEPI 76

Qy 61 RLQVPVEAVNMSLGLSLAATTPPOLACGPTVHOTCSNTYVKGCLFGLFSGNLROQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPOLACGPTVHOTCSNTYVKGCLFGLFSGNLROQPOK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMBQKSKTLFSLMQYSSEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMBQKSKTLFSLMQYSSEF 196
Qy 181 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPRDRHVFQV 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPRDRHVFQV 316
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSFEHMSQEGFSAITSNGLPSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQTGSSSFEHMSQEGFSAITSNGLPSTVGSYDWAG 376
Qy 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYEQTRGGQVSVCP 496
Qy 481 PRGQARWOCDAVLGYEQOPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
Db 497 PRGQARWOCDAVLGYEQOPNGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFCQSLGGQDLTMDGLVDLTGVAQGHVLLLRQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFCQSLGGQDLTMDGLVDLTGVAQGHVLLLRQ 616
Qy 601 PVLRYKAIEMEPREVARNVFCNDQVVKGEVRVCLHVOKSTROPRLRGOIQSVVT 660
Db 617 PVLRYKAIEMEPREVARNVFCNDQVVKGEVRVCLHVOKSTROPRLRGOIQSVVT 676
Qy 661 YDLALDSDGRPSRAVFNKSTRQTVLGTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSDGRPSRAVFNKSTRQTVLGTQTCETLKLQLPNCIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLADARLFTALPFPFKNCNDNICODDLSITPFSMSLCLVVG 780
Db 737 SLVGTPLSAFNLRPVLADARLFTALPFPFKNCNDNICODDLSITPFSMSLCLVVG 796
Qy 781 GPRENVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSRMLACESASSTEV 840
Db 797 GPRENVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSRMLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFTFDVDSKASLGNKLLKANVTSENMPRNTKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFTFDVDSKASLGNKLLKANVTSENMPRNTKTEF 916
Qy 901 QLELPVKYAVVMVTSYKYNFTASNTSRVQHQYQVSNLQQRSLPISLFLVPV 960
Db 917 QLELPVKYAVVMVTSYKYNFTASNTSRVQHQYQVSNLQQRSLPISLFLVPV 976
Qy 961 RLNOTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRTQCDIP 1020
Db 977 RLNOTVIMDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRTQCDIP 1036
Qy 1021 FFGIOBEFNATLKGNSLFDWYKTSNHLIIYSTAIELFNDVSVFTLLPQCGAFVRSQTEF 1080
Db 1037 FFGIOBEFNATLKGNSLFDWYKTSNHLIIYSTAIELFNDVSVFTLLPQCGAFVRSQTEF 1096
Qy 1081 KVEPPEVPNPLIIVGSSVGGLLLLALITAALYKLGFFKRYKQKDMWSEGGPGABPQ 1137
Db 1097 KVEPPEVPNPLIIVGSSVGGLLLLALITAALYKLGFFKRYKQKDMWSEGGPGABPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286.889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173.497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: William Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3
Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFGQSVVOLQSGRVVVGAPQEIIVAAANORGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFGQSVVOLQSGRVVVGAPQEIIVAAANORGLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTPPOLACGPTVHOTCSNTYVKGCLFGLFSGNLROQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPOLACGPTVHOTCSNTYVKGCLFGLFSGNLROQPOK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMBQKSKTLFSLMQYSSEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMBQKSKTLFSLMQYSSEF 196
Qy 181 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPRDRHVFQV 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPRDRHVFQV 316
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSFEHMSQEGFSAITSNGLPSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQTGSSSFEHMSQEGFSAITSNGLPSTVGSYDWAG 376

QY 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVCPPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVCPPL 496
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVLYF 540
DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGQDGLTMDGLVLDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGQDGLTMDGLVLDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMFNPREVARNVPECDQVVKGEKAGEVRVCLHVQKSTRDRRLREGIQSVVT 660
DB 617 PVLRVKAIMFNPREVARNVPECDQVVKGEKAGEVRVCLHVQKSTRDRRLREGIQSVVT 676
QY 661 YDLALDSGRHSRAVFNKSTRQVGLTQTCETLKLQPNCTIEDPVSPIVLRNLF 720
DB 677 YDLALDSGRHSRAVFNKSTRQVGLTQTCETLKLQPNCTIEDPVSPIVLRNLF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAQRLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNLRPVLAEDAQRLFTALFPFKKNCNDNIQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLVSRKVSTLQNRQSRQSWRLACASSTEV 840
DB 797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLVSRKVSTLQNRQSRQSWRLACASSTEV 856
QY 841 SGALKSTSCSINHPPIPFENSEVFNITFDVDSKASLGKLLKLVNTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPPIPFENSEVFNITFDVDSKASLGKLLKLVNTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHQVSNLQSRSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHQVSNLQSRSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 FFGQEBFNATLKNLSFDVYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGFVRSQTET 1080
DB 1037 FFGQEBFNATLKNLSFDVYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGFVRSQTET 1096
QY 1081 KVEPFEVNPPLPLIVGSSVGLLALLITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1137
DB 1097 KVEPFEVNPPLPLIVGSSVGLLALLITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1153

RESULT 3

US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,618

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/173,497

FILING DATE: 23-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/362,652

FILING DATE: 21-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659

REFERENCE/DOCKET NUMBER: 27866/32797

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

TELEFAX: 25-3856

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1153 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-485-618-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDENAMTFQENARGFGQSVVQLQSGSRVVVGAPOEIVAAANRGSLYQCDYSTGSCPEI 60

DB 17 FNLDENAMTFQENARGFGQSVVQLQSGSRVVVGAPOEIVAAANRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQOPQK 120

DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHPFRMKFVSTVMEOLKSKTLPSLMQYSEEF 180

DB 137 FPEALRGCPQEDSDIAFLIDGSGIIPHPFRMKFVSTVMEOLKSKTLPSLMQYSEEF 196

QY 181 RIHFTFKFQNNPNRSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240

DB 197 RIHFTFKFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGVEDVITPEADREGVIRYVIGVGDAFSEKSRQBELNTIASKPPRDHVPQVN 300

DB 257 TDGEKFGDPLGVEDVITPEADREGVIRYVIGVGDAFSEKSRQBELNTIASKPPRDHVPQVN 316

QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATSGNPLLSITVGSYDWAG 360

DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATSGNPLLSITVGSYDWAG 376

QY 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR 420

DB 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVCPPL 480

DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYHYEQTRGGQSVCPPL 496

QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVLYF 540

DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVLYF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGQDGLTMDGLVLDLTVGAQGHVLLRSQ 600

Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGDLTMDGLVDTLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHRAVFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHRAVFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFAGNLRLPVLAEADQRLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFAGNLRLPVLAEADQRLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFFPLDLSYRKVSTLQORSQSRWFLACASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFFPLDLSYRKVSTLQORSQSRWFLACASSTEV 856
QY 841 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTET 900
Db 857 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTET 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVLPLVP 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVLPLVP 976
QY 961 RLNQTVINDRPOVTFSENISSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNQTVINDRPOVTFSENISSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036
QY 1021 FFGIOEFENATLKNLSFQWYIKTSHNLLIVSTAEILFNDVFTLLPGGAPVRSOTET 1080
Db 1037 FFGIOEFENATLKNLSFQWYIKTSHNLLIVSTAEILFNDVFTLLPGGAPVRSOTET 1096
QY 1081 KVEPFVFPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1137
Db 1097 KVEPFVFPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABEQ 1153

RESULT 4

US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLAGCPTVHQTCSNTYVVKGLCLFGLSNLRQQPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLAGCPTVHQTCSNTYVVKGLCLFGLSNLRQQPQK 136
QY 121 PPEALRGCPQEDSDTAFLIDSGSII PHDFRMKEFVSTVMEQLKSKTKLFSLMQYSEEF 180
Db 137 PPEALRGCPQEDSDTAFLIDSGSII PHDFRMKEFVSTVMEQLKSKTKLFSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNTTNGARKNAFKILVVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNTTNGARKNAFKILVVI 256
QY 241 TDGEFGDPLGVYEDVIPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDHVFQVN 300
Db 257 TDGEFGDPLGVYEDVIPEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDHVFQVN 316
QY 301 NPEALKTIONLREKI FAIEGTQTGSSSFHEHENSEGFSAAITNSGPLLSTVGSYDWAQ 360
Db 317 NPEALKTIONLREKI FAIEGTQTGSSSFHEHENSEGFSAAITNSGPLLSTVGSYDWAQ 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVSCPL 496
QY 481 PRGQBARWQCDVLYGEOGQPHGRFGAALTVLGDVNGDKLTVDVAIGAPGEEDNRGAVLYF 540
Db 497 PRGQBARWQCDVLYGEOGQPHGRFGAALTVLGDVNGDKLTVDVAIGAPGEEDNRGAVLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGDLTMDGLVDTLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGDLTMDGLVDTLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHRAVFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHRAVFNETKSTRRTQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFAGNLRLPVLAEADQRLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFAGNLRLPVLAEADQRLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFFPLDLSYRKVSTLQORSQSRWFLACASSTEV 840

797 GPRFNVTVVRNDGSDSVRTQVTFPPFLDLSVRKYSTLQNRQSRWRLACESASSTEV 856
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 900
857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 916
901 QLELPVKYAYVMVTVSHGVSTKYLNFTASNTSRVVMQHQYQVSNLQORSLPISLVLVPV 960
917 QLELPVKYAYVMVTVSHGVSTKYLNFTASNTSRVVMQHQYQVSNLQORSLPISLVLVPV 976
961 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020
977 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036
1021 FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSFTLLPGOGAFVRQOTET 1080
1037 FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSFTLLPGOGAFVRQOTET 1096
1081 KVPEFVNPPLPIVGVSSVGLLLALLITAAALYKLGFFKQYKDMVSEGGPPGAEPQ 1137
1097 KVPEFVNPPLPIVGVSSVGLLLALLITAAALYKLGFFKQYKDMVSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARGFQGSVVQLOGSRVVVVGAPQEIIVAANQORGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFQGSVVQLOGSRVVVVGAPQEIIVAANQORGLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVVGKLCFLFGNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQCSNTYVVGKLCFLFGNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTMEOQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEVSTMEOQLKSKTFLSLMOYSEEF 196
Qy 181 RHFTFKFQNNPNRSLVKPITOLLGRHTATGVRKVIKRELLNITNGARKNAFKILVI 240
Db 197 RHFTFKFQNNPNRSLVKPITOLLGRHTATGVRKVIKRELLNITNGARKNAFKILVI 256
Qy 241 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVTPADREGVIRYVIGVDAPRSEKSRQBLNTIASKPPRDHVFQVN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATITNSGPLLSTVGSYDAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATITNSGPLLSTVGSYDAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNAYLYAAAILRNVRQSLVLCAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNAYLYAAAILRNVRQSLVLCAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQGVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQGVSVCP 496
Qy 481 PRQARARWOCDAVLYGEOQPMGRFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQARARWOCDAVLYGEOQPMGRFCAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
Qy 601 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRQGIQSVVT 660
Db 617 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRQGIQSVVT 676
Qy 661 YDLALDSGRPHSRAVENETKNSRROTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRAVENETKNSRROTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFFPKNCNDNI CODDLSITFSFMSLCLVWG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFFPKNCNDNI CODDLSITFSFMSLCLVWG 796
Qy 781 GPREFNVTVVRNDGEDSVYRTQVTFPPFLDLSVRKYSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVYRTQVTFPPFLDLSVRKYSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAYVMVTVSHGVSTKYLNFTASNTSRVVMQHQYQVSNLQORSLPISLVLVPV 960
Db 917 QLELPVKYAYVMVTVSHGVSTKYLNFTASNTSRVVMQHQYQVSNLQORSLPISLVLVPV 976
Qy 961 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020
Db 977 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036

QY 1021 FFGIOEEFNATLKGNSLSPDWNKTSNHNHLLIVSTAEILFNDVSFTLLPQGAFAVRSOTET 1080
DB 1037 FFGIOEEFNATLKGNSLSPDWNKTSNHNHLLIVSTAEILFNDVSFTLLPQGAFAVRSOTET 1096
QY 1081 KVEPEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPQ 1137
DB 1097 KVEPEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPQ 1153

RESULT 6

US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marehall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,859
; REFERENCE/DOCKET NUMBER: 127866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 76
QY 61 RLOVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOPK 120
DB 77 RLOVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOPK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKSKSLFSLMQYSEEF 180

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQLKSKSLFSLMQYSEEF 196
QY 181 RIHFTPEFONNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGAKNAFKILIVI 240
DB 197 RIHFTPEFONNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGAKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNITIASKPPRDHVFQVN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNITIASKPPRDHVFQVN 316
QY 301 NFEALKTIONLREKIPAIETGOTGSSSFEHEMSQEGPSAAITNSGNPLISTVGSYDAG 360
DB 317 NFEALKTIONLREKIPAIETGOTGSSSFEHEMSQEGPSAAITNSGNPLISTVGSYDAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLVGAATAIILNRVQSLVGLGAPRQHIHGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLVGAATAIILNRVQSLVGLGAPRQHIHGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTTRGGQSVCP 496
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 540
DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLLRQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLLRQ 616
QY 601 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTDRDLREGQIQSVVT 660
DB 617 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTDRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRTQVLTCTETLKLQPNCTEDPVSPIVLRNLF 720
DB 677 YDLALDSGRPHSRVAFNETKSTRTQVLTCTETLKLQPNCTEDPVSPIVLRNLF 736
QY 721 SLVGTPLSAFONLRPVLAEADQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFONLRPVLAEADQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQORSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNSIAYCQRIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNSIAYCQRIQCDIP 1036
QY 1021 PFGIOEEFNATLKGNSLSPDWNKTSNHNHLLIVSTAEILFNDVSFTLLPQGAFAVRSOTET 1080
DB 1037 PFGIOEEFNATLKGNSLSPDWNKTSNHNHLLIVSTAEILFNDVSFTLLPQGAFAVRSOTET 1096
QY 1081 KVEPEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPQ 1137
DB 1097 KVEPEVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKMMSEGGPPGAEPQ 1153

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGSVVOLGSRVVVGAPQEIIVANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGSVVOLGSRVVVGAPQEIIVANORGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSGLSLAATSPQILLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSGLSLAATSPQILLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136

Qy 121 FPEALRCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVWEQLKSKTLPFLMQYSEEF 180
Db 137 FPEALRCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVWEQLKSKTLPFLMQYSEEF 196

Qy 181 RIHFTKFEQNNPNSRLVPIPTQLLGRTHATGATGVRKVIKRLNITNGARKNAFKILIVI 240
Db 197 RIHFTKFEQNNPNSRLVPIPTQLLGRTHATGATGVRKVIKRLNITNGARKNAFKILIVI 256

Qy 241 TDEKFGDPLGYEDVIEADREGVIRVIGVGDAFRSEKSRQELNITASKPPRDHVFQVN 300
Db 257 TDEKFGDPLGYEDVIEADREGVIRVIGVGDAFRSEKSRQELNITASKPPRDHVFQVN 316

Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYIYETRGQSVSVCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYIYETRGQSVSVCPL 496

Qy 481 PRGORARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYLF 540
Db 497 PRGORARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYLF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGODLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGODLTMDGLVDLTGVAQGHVLLRSQ 616

Qy 601 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 617 PVLRVKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRVAFNETKNSRRROTQVLGLTQTCTETLKLQLPNCIEDPVPSPVLRNLF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRRROTQVLGLTQTCTETLKLQLPNCIEDPVPSPVLRNLF 736

Qy 721 SLVGTPLSAFAGNLRLPVLAEQAORLFTALFPFPEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFAGNLRLPVLAEQAORLFTALFPFPEKNCGNDNICQDDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPFDLDSYRKVSTLONQSRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPFDLDSYRKVSTLONQSRQSRWLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIFPENSEVENTITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVENTITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 916

Qy 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVNMHQYQVSNLQORSIPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVNMHQYQVSNLQORSIPISLVFLVPV 976

Qy 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNVNCSTAVCQRIQCDIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNVNCSTAVCQRIQCDIP 1036

Qy 1021 FFGIOEEFNATLKGNLSDFWYIKTSHNHLIIVSTABILFNDVSFTLLPGOGAFVRSOTET 1080
Db 1037 FFGIOEEFNATLKGNLSDFWYIKTSHNHLIIVSTABILFNDVSFTLLPGOGAFVRSOTET 1096

Qy 1081 KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMMSSEGGPPGAEPQ 1137
Db 1097 KVEPFEPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMMSSEGGPPGAEPQ 1153

RESULT 8

US-09-193-043-3
Sequence 3, Application US/09193043
Patent No. 6251395
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6251395el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 99.9%; Score 5868; DB 3; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGLSYOCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGLSYOCDYSTGSCPEI 76

Qy 61 RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRRQPOK 136

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMBOLKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMBOLKSKTFLSLMQYSEEF 196

Qy 181 RIHFTKFEQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTKFEQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQV 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQV 316

Qy 301 NFEALKTIONLREKIFAIEGTQSGSSSFEHMSQEGFSAAITSGPLLSVGSYDWAG 360
Db 317 NFEALKTIONLREKIFAIEGTQSGSSSFEHMSQEGFSAAITSGPLLSVGSYDWAG 376

Qy 361 GVPLYTSKEKSTFINNTRVDSMDNAYLGAAIILRNVRQSLVGLGAPYQHIGLVAMFR 420
Db 377 GVPLYTSKEKSTFINNTRVDSMDNAYLGAAIILRNVRQSLVGLGAPYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLI GAPHYEYQTRGGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLI GAPHYEYQTRGGQSVCP 496

Qy 481 PRGQARWQCDVLYGEOGPWGRFCAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEOGPWGRFCAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

Qy 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGODLTMDGLVDLTGCAQGHVLLLRQ 600
Db 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGODLTMDGLVDLTGCAQGHVLLLRQ 616

Qy 601 PVLRVKAIMEFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVYT 660
Db 617 PVLRVKAIMEFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVYT 676

Qy 661 YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETLKLQPNICIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETLKLQPNICIEDPVSPVILRLNF 736

Qy 721 SLVGTPLSAGNLRPVLAEDAQRFLTALPFPEKNCNDNITCDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAGNLRPVLAEDAQRFLTALPFPEKNCNDNITCDDLSITFSFMSLDCLVVG 796

Qy 781 GPREFNVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWSRLACESASSTEV 840
Db 797 GPREFNVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWSRLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGKLLKLLKANTYSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGKLLKLLKANTYSENMPRTNKTFF 916

Qy 901 QLELPVKYAVVMVTSHGVSSTKYLNFNTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 960

Db 917 QLELPVKYAVVMVTSHGVSSTKYLNFNTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 976

Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCORIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCORIQCDIP 1036

Qy 1021 FFGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
Db 1037 FFGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1096

Qy 1081 KVEPFEVNPPLPLIVGSSVGGLLLLALITAAALYKGLGFFKQYKDMMEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLPLIVGSSVGGLLLLALITAAALYKGLGFFKQYKDMMEGGPPGAEPQ 1153

RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 99.9%; Score 5868; DB 4; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGLSYOCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGLSYOCDYSTGSCPEI 76

Qy 61 RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRRQPOK 136

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMBOLKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMBOLKSKTFLSLMQYSEEF 196

Qy 181 RIHFTKFEQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTKFEQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQV 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQV 316

Qy 301 NFEALKTIONLREKIFAIEGTQSGSSSFEHMSQEGFSAAITSGPLLSVGSYDWAG 360

```
Db 317 NFEALKTIONQREKI FAIEGTQTGSSSSFEHEMSQEGFSAATSNQPLLSLTVGSYDWAQ 376
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALILNRRVQSLVLAGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALILNRRVQSLVLAGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYLYF 540
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYLYF 556
Qy 541 HGTSGSGISPSHSORIASGKSLSPLOYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKSLSPLOYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMEFNPREVARNPECDNDQVVKGEAGEVRVCLHVQKSTRDRLRREGIQSVVT 660
Db 617 PVLRVKAIMEFNPREVARNPECDNDQVVKGEAGEVRVCLHVQKSTRDRLRREGIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVGLTQTCETKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFFSMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFFSMSLDCLVWG 796
Qy 781 GPREFNVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSRWLACASSTEV 840
Db 797 GPREFNVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSRWLACASSTEV 856
Qy 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVVMVTSYKLNFTASENTSRVMOHQVSNLQORSLSPLISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSYKLNFTASENTSRVMOHQVSNLQORSLSPLISLVFLVPV 976
Qy 961 RLNTQVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDIP 1020
Db 977 RLNTQVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDIP 1036
Qy 1021 FFGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSVFTLLPGQCAFVRSOTET 1080
Db 1037 FFGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSVFTLLPGQCAFVRSOTET 1096
Qy 1081 KVEPFEVNPPLPLIVGSSVGLLALLITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLPLIVGSSVGLLALLITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1153
```

RESULT 10

```
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnsaut, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-476-062A-43
```

```
Query Match 99.4%; Score 5837.5; DB 2; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136
Qy 121 FPALAGCPQEDSDIAFLIDGSGSIIPHDPRRMEKFVSTVMEQLKSKTLFSLMOYSEEF 180
Db 137 FPALAGCPQEDSDIAFLIDGSGSIIPHDPRRMEKFVSTVMEQLKSKTLFSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLVKPIQTQLGRTHATGVRKVIRELNIITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPIQTQLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVITPEADREGVIRYVIGVGDAFRSEKSRQBELNTIASKPPRDHVPQVN 300
Db 257 TDGEKFGDPLGYEDVITPEADREGVIRYVIGVGDAFRSEKSRQBELNTIASKPPRDHVPQVN 316
Qy 301 NFEALKTIONQREKI FAIEGTQTGSSSSFEHEMSQEGFSAATSNQPLLSLTVGSYDWAQ 360
Db 317 NFEALKTIONQREKI FAIEGTQTGSSSSFEHEMSQEGFSAATSNQPLLSLTVGSYDWAQ 376
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALILNRRVQSLVLAGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALILNRRVQSLVLAGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYLYF 540
Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAYLYF 555
Qy 541 HGTSGSGISPSHSORIASGKSLSPLOYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
```


Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPIISLVFLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPIISLVFLVPV 975
Qy 961 RLNOTVWDPRQVTFSENLSTCHTKERLPSSHDFLAELRKAPVWNCISIAVCORIQCIP 1020
Db 976 RLNOTVWDPRQVTFSENLSTCHTKERLPSSHDFLAELRKAPVWNCISIAVCORIQCIP 1035
Qy 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
Db 1036 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGPPGAEPO 1137
Db 1096 KVEPFEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGPPGAEPO 1152

RESULT 12
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION NUMBER: US/08/78,871
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO:2
; LENGTH: 1152

5424399-2
Query Match 99.4%; Score 5837.5; DB 6; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARFGSGSVVQLQGRVWVWGAPOIIVAAQORGSILYQCDYTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGSGSVVQLQGRVWVWGAPOIIVAAQORGSILYQCDYTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHOTCSENTYKVLGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHOTCSENTYKVLGCLFLFGSNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPDFRRMKFVSTVMEQLKSKTILFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPDFRRMKFVSTVMEQLKSKTILFSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPNSRLVPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 240
Db 197 RIHFTFKFQNNPNSRLVPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 256
Qy 241 TDGKFGDPLGYEDVIEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDHVFQVN 300
Db 257 TDGKFGDPLGYEDVIEADREGVIRYVIGVGDAPFRSEKSRQELNTIASKPPRDHVFQVN 316
Qy 301 NFEALKTIQOLREKIFAIEGTOTGSSSSFEHEMSQSGFSAATISNGPLLSSTVGSYDWAG 360
Db 317 NFEALKTIQOLREKIFAIEGTOTGSSSSFEHEMSQSGFSAATISNGPLLSSTVGSYDWAG 376
Qy 361 GVFLYTSKESKSTFNNMTRVSDMNDAYLGYAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFNNMTRVSDMNDAYLGYAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMEESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGGOVSCPL 480
Db 437 QNTGMEESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGGOVSCPL 496
Qy 481 PRGORARWQCDVAVLYGQGGPWRFGAALTVLGVDNGDKLTDVAIGAPEGEDNRGAVYLF 540

Db 497 PRG-RARWQCDVAVLYGQGGPWRFGAALTVLGVDNGDKLTDVAIGAPEGEDNRGAVYLF 555
Qy 541 HGTSSGSISSPHSQRSIAGSKLSPRLQYFQOSLSGGQDLTMDGLVDLTGGAQGHVLLLRSQ 600
Db 556 HGTSSGSISSPHSQRSIAGSKLSPRLQYFQOSLSGGQDLTMDGLVDLTGGAQGHVLLLRSQ 615
Qy 601 PVLRVKAIEMFNPREVARNVFECDQVVKAGKAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 616 PVLRVKAIEMFNPREVARNVFECDQVVKAGKAGEVRVCLHVQKSTRDLRREGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKSTRRTQVLTGTTCTETIKLQLPNCIEDPVSPIVLRNLF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRRTQVLTGTTCTETIKLQLPNCIEDPVSPIVLRNLF 735
Qy 721 SLVGTPLSLAFGNLRPVLAEDAORLFTALFPFKNCGNDNICDDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSLAFGNLRPVLAEDAORLFTALFPFKNCGNDNICDDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 796 GPREFNVTVRNDEGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANTVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANTVTSENNMPTNKTEF 915
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPIISLVFLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSPIISLVFLVPV 975
Qy 961 RLNOTVWDPRQVTFSENLSTCHTKERLPSSHDFLAELRKAPVWNCISIAVCORIQCIP 1020
Db 976 RLNOTVWDPRQVTFSENLSTCHTKERLPSSHDFLAELRKAPVWNCISIAVCORIQCIP 1035
Qy 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
Db 1036 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGPPGAEPO 1137
Db 1096 KVEPFEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGPPGAEPO 1152

RESULT 13
US-08-476-062A-44
; Sequence 44, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842

[illegible]

Db 318 DFDALDKIQNLKEKIFAIEGTETITSSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG 377
QY 361 GVFLVTSKEKSTFINMTVDSDMDAYLGAAAIILNRNVOSLVLGAPRYOHIGLVAMFR 420
Db 378 GAFLYPPNPNSTFINNSQENVDMRDSYIGYSTELAKWGVOSLVLGAPRYOHIGKAVIFI 437
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTIGSYFGASLCSVDVDTGSTDVLVIGAPHYYEQTRGGQVSVCP 497
QY 481 PRGQARWOCDAVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEBEDNRGAVYLF 540
Db 498 PRGWR-RWMCDAVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVVI GAPGEBEENRGAVYLF 556
QY 541 HCTSGSGISPSHSQRIAGSKLPRIQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLGPSISPSHSQRIAGSKLSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616
QY 601 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLWGVSNQFIPAIIPRSAFECEQVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHRAVFNETHNSTRQTVLGTQTCETIKLQLPNCIEDPVPVILRLNF 720
Db 677 LDALAPGRLSPRAIFQETKRSLSRVRVLGKAHCENFNLLSPSCVEDSVIPIILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAORLEFALPPEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 TLVGPPLLAFLNRPMLAALARYFTASLPFEKNCGADHICODNLGISFSFPLGKSLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 SNLELNAEVMVWNGEDSVGTITFHPAGLSYRYVAEGQKQQLRSLHLTC--CSAPVG 854
QY 841 SGALKSTCSINHPIPPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 855 SQGTWSTSCRINHPIPRGCAQITFLATFDVSPKAVGLDRLLLIANVSENNIPTSKTIF 914
QY 901 QLELPVKYAVYVWVTSKYNFTAS-ENTSRVMQHOYOVSNLQORSLPISLVELVP 959
Db 915 QLELPVKYAVYVWVTSKYNFTAS-ENTSRVMQHOYOVSNLQORSLPISLVELVP 974
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI 1019
Db 975 VELNQEAVMWDEVSHQVPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGLRFRCDV 1034
QY 1020 PFFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDVSVFTLLPGQAFVRSQTE 1079
Db 1035 PGSFVQEEELDFTLKGNLSFGWVRQILQKKVSVVVAEIIFTSVYSQLPQOEFMRAQTI 1094
QY 1080 TKVEPPEVPNPLIIVSSVGLLILALITAAVLYKLGFEKROYKDMWSE 1128
Db 1095 TVLEKYKVNPIPLIIVSSIGGLLILALITAVLYKVGFFKROYKEMWEE 1143

Search completed: November 25, 2003, 14:23:24
Job time : 16.5618 secs